ARL 65-116, PART I
JUNE 1965

Aerospace Research Laboratories

DESIGN, TESTING AND ESTIMATION IN COMPLEX EXPERIMENTATION

I. EXPANSIBLE AND CONTRACTIBLE FACTORIAL DESIGNS AND THE APPLICATION OF LINEAR PROGRAMMING TO COMBINATORIAL PROBLEMS

STEVE WEBB
ROCKETDYNE
A DIVISION OF NORTH AMERICAN AVIATION, INC.
CANOGA PARK, CALIFORNIA

OFFICE OF AEROSPACE RESEARCH
United States Air Force

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

- - - - - - - - - -

Qualified requesters may obtain copies of this report from the Defense Documentation Center, (DDC), Cameron Station, Alexandria, Virginia.

- - - - - - - - - -

This report has been released to the Office of Technical Services, U. S. Department of Commerce, Washington 25, D. C. for sale to the general public.

- - - - - - - - - -

Copies of ARL Technical Documentary Reports should not be returned to Aerospace Research Laboratories unless return is required by security considerations, contractual obligations or notices on a specified document.

700 - July 1965 - 448-48-1034

BLANK PAGE

ARL 65-116
PART I

DESIGN, TESTING AND ESTIMATION IN COMPLEX EXPERIMENTATION

I. EXPANSIBLE AND CONTRACTIBLE FACTORIAL DESIGNS AND THE APPLICATION OF LINEAR PROGRAMMING TO COMBINATORIAL PROBLEMS

STEVE WEBB

ROCKETDYNE
A DIVISION OF NORTH AMERICAN AVIATION, INC.
CANOGA PARK, CALIFORNIA

JUNE 1965

Contract No. AF 33(616)-7372
Project No. 7071
Task No. 7071-01

AEROSPACE RESEARCH LABORATORIES
OFFICE OF AEROSPACE RESEARCH
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

This is part one of a final report on work sponsored by the Aerospace Research Laboratories, Office of Aerospace Research, United States Air Force, under Contract AF 33(616)-7372, Project 7071, Task 7071-01. The technical effort, conducted at Rocketdyne in the Research Department, was monitored by Dr. H. Leon Harter of the Aerospace Research Laboratories.

The initial motivation of the overall research program was to develop methods for finding optimum experimental designs, taking into account cost as well as variance properties. Loss functions which included quadratic and linear cost terms were used for designs based on simple regression models. Explicit minimizations of this loss function were obtained in certain simple cases (e.g., strictly linear multivariate model) and some interesting general nonoptimality properties of orthogonal designs were discovered. Treatment of more general cases, however, met with excessive computational difficulty and for this reason the scope of the original program was expanded to include the use of Bayesian decision theory and finally more recently the application of the methods of linear programming. New tasks were subsequently added in the areas of expansible and contractible designs, multivariate quantal response problems, and most recently estimation questions in reliability growth models.

The research on loss functions has been described in detail in previous ARL Technical Documentary Reports (ARL 62-373, ARL 63-107) and no further mention will be made of it here. Results obtained in the other investigations during the past two years (ending February, 1965) are reviewed in the three parts of this report. Part two deals with formulations of reliability growth models and statistical

estimation of parameters of the stochastic processes resulting from such models. Part three involves the design and analysis of sensitivity experiments in which there are one or more stimulus variables.

Portions of the work reported in part one (Appendices A, B, and E) were based on dissertation research by the author while he was a graduate student at the University of Chicago. The author acknowledges the assistance of Mr. George Uglean on various parts of this research.

ABSTRACT

This report describes results of research on factorial designs during a two-year period ending in February, 1965. These include (a) characterizations of orthogonal and two classes of non-orthogonal designs as solutions to linear constraints, (b) optimality properties of orthogonal designs, (c) development of a general class of non-orthogonal sequential factorial designs, (d) results on certain families of 2^n designs, and (e) description of a special-purpose linear programming computer routine for combinatorial problems in experimental design.

TABLE OF CONTENTS

SUMMARY

Results obtained along two lines of study are given in this report. The first is concerned with a special class of sequential incomplete factorial designs, termed "expansible and contractible". The study of this class was motivated by the need for designs which do not lose their desirable properties if they are stopped prematurely or if new factors are added after the experiment has been initiated. While any design can be ordered so that the factors are introduced sequentially, not all designs have the contractibility property, namely that truncated designs retain relatively high efficiency. In particular, orthogonal designs do not usually provide convenient expansible series of designs. The class of one-at-a-time designs is a convenient expansible series, but the variance efficiencies of such designs are poor. The general class of permutation-invariant designs is defined and discussed in detail, and it is shown that many expansible series can be constructed within this class. The variances and efficiencies for many resolution-five 2^n designs are given, and the best expansible series are indicated.

The second investigation involves the use of linear programming for finding incomplete factorial designs. Details of the simplex method are reviewed, and a special-purpose linear-programming computer routine for finding factorial designs is described. It is shown how the special features of the problems, such as the size and condition of the basis matrices and the degeneracy of the solutions, are exploited in the computer routine. It has been used successfully to derive several known designs which involve constraint matrices of moderate size. On larger problems, computational difficulties and non-integral solutions have precluded obtaining new designs.

Manuscript released by the author May 1965 for publication as an ARL Technical Report.

This report has as appendices five papers which had previously been issued only as internal reports at Rocketdyne. Appendix A deals with the formulation of combinatorial problems in experimental design as linear programming problems. Appendix B contains a survey of non-orthogonal designs, introduces the classes of permutation-invariant and clumpwise-orthogonal designs, and gives linear-programming characterizations of these classes. Appendix C is a paper on optimality properties of orthogonal designs which was presented at the Chicago meeting of the American Statistical Association in December of 1964. Appendix D is an investigation of the class of one-at-a-time designs. Appendix E contains results on incomplete factorial designs originally reported at the Urbana meeting of the Institute of Mathematical Statistics in November of 1961.

EXPANSIBLE AND CONTRACTIBLE DESIGNS

INTRODUCTION

If an ordinary incomplete factorial design is halted prematurely, most of the information is lost. In industrial research and development programs, there are many reasons why a planned experiment may be stopped, including budgetary changes, failure of test equipment, pressure for results, or a shift in experimental emphasis as it becomes obvious that a factor should be included or excluded. The uncertainty of the future of a program often leads an experimenter to resist the apparent rigidity of a formal statistical test plan. The concept of designs which introduce the factors sequentially, which we have chosen to call expansible and contractible designs, has been developed to correct this deficiency in the designs currently available in the literature.

A contractible design is an ordered specification of treatment combinations at which runs are to be made which has the property that if the experiment is halted prematurely, quantitative inferences about the effects of some of the factors can still be made. Contractibility can be achieved by varying only a few of the factors, preferably those thought beforehand to have the largest effects, in the early part of the experiment. If the factors which are held fixed do not interact with those which are varied, then the effects of the factors which have been varied can be estimated without bias. The unestimated effects are, however, aliased with the estimates of the mean.

Any design is expansible since any design can be repeated with a previously fixed factor at another level. It is useful to apply the term, however, to series of designs any one of which can be contracted to the next lower or expanded to the next higher design in the series.

As an example of these ideas we consider a simple expansible and contractible series for up to three two-level factors. For one factor the design is simply

0

1 ,

where 0 represents the low level and 1 the high level of the single factor. Now suppose a second factor, which was at its low level during the first two runs, is varied. The design is

0 0

1 0

0 1 ,

of which the first two runs have already been done. A third factor (at its low level for the first three runs) may now be added in two additional runs, so that the complete design is

0 0 0

1 0 0

0 1 0

0 0 1

1 1 1

In unpublished work this design has been shown by the methods of Appendix E to be the five-run design for which the variance of each estimate is minimized. After the first two tests, note that the estimate of the mean is aliased with the main effects of the second and third factors, but the estimate of the main effect of the first factor is not aliased. Similarly, after three runs the main effects of the first and second factors can be estimated while the effect of the third is

aliased with the mean. After four runs the effects of all three factors can be estimated, but the addition of the fifth test improves the efficiency of the estimates considerably. Note that the last four tests comprise a half-replicate of a 2^3, which is the standard design to use for three noninteractive factors. This half-replicate is not contractible, however, and no estimates can be calculated until after all four responses have been observed.

Below is described the general theory applicable to expansible and contractible designs for 2^n experiments. The use of expansible series of orthogonal designs is discussed as well as one-at-a-time designs. Finally, the properties of a general class of such designs, permutation-invariant designs of resolution 5, are derived in detail.

EXPANSIBLE AND CONTRACTIBLE DESIGNS FOR TWO-LEVEL FACTORS

General Theory

The treatment combinations of any incomplete 2^n design can be arranged in such an order that the factors are introduced sequentially. Here and in the sequel, a factor is considered to be introduced at that run in which the factor first appears at its high level. If one considers the treatment combinations as n-bit binary numbers, an arrangement with this property may be obtained by placing the treatment combinations in increasing numerical order. If the factors are arranged in order of increasing a priori importance (increasing magnitude of effect), then the order of introduction is the inverse of the order of importance. A design ordered in this way may or may not be contractible depending on whether or not the subdesigns are singular. For example, the subdesigns of the half-replicate of a 2^3

0 0 1

0 1 0

1 0 0

1 1 1

are singular.

For 2^n designs it can be shown that the variance of each estimate is minimized if and only if the design is orthogonal. It is natural to inquire therefore, whether orthogonal designs are contractible and expansible. Since in an orthogonal incomplete 2^n design the high level appears in half the tests, an orthogonal design would necessarily be contracted to half its number of runs or expanded to twice its number of runs by the deletion or addition of a factor. Therefore, if there is pressure to hold the number of tests to a minimum, it will usually not be useful to restrict attention to orthogonal designs.

There are several exceptions to the general rule that the contractibility property is not satisfied by the orthogonal design which has the smallest number of runs for a given experiment. The only exception for resolution 3 designs (i.e., those for estimating only main effects) is the trivial case of two factors, for which the minimal four-run design can be contracted to the minimal two-run one-factor design. There are several exceptions for resolution 5 designs (those for estimating main effects and two-factor interactions). These cases are: expanding from one up through four factors; expanding from five up through seven factors; and expanding from eight to nine factors. In all these cases the smallest orthogonal design containing $i+1$ factors contains twice as many runs as the smallest for i.

The optimality property which leads one to select orthogonal designs is based on variance considerations. Such considerations may be subordinate to others in research and development programs, in which it often happens that the cost per test is large and the error of an individual observation is smaller in magnitude than the effects of interest. In the extreme case where the error variance is zero, the only relevant criterion for a design is its "degree-of-freedom" efficiency, defined as the ratio of the number of parameters to the number of runs. In evaluating 2^n designs one may take an intermediate point of view and consider the efficacy[1] of designs, which is the product of the degree-of-freedom efficiency and the prediction index, defined in another report [3]. The prediction index for 2^n designs reduces to the reciprocal of the product of the number of runs and the average variance of an estimate. The index has its maximum value of unity for orthogonal (equal-frequency) designs.

One more concept which is useful in developing expansible series of designs is the "guaranteed minimum". It may often be assumed that the experiment will not be stopped before a certain number of runs have been made, which number we refer to as the guaranteed minimum. It is then not necessary to require that the design be contractible beyond this minimum. If a design has no guaranteed minimum associated with it, then it is said to be fully contractible.

One-At-A-Time Designs

A 2^n design containing $n+1$ runs in which the i-th factor is introduced at the $(i+1)$-st run is called a one-at-a-time design. Such designs are of resolution 3, are fully contractible, and are perfectly efficient in the degree-of-freedom sense. Their variance properties are, however, known to be poor. In exploratory

[1] This term and the concept were suggested to the author by Cuthbert Daniel.

research with low error their degree-of-freedom efficiency and contractibility may still make them useful designs. In another paper, which is included as Appendix D, it was shown that $\frac{1}{2}\sigma^2$ is a lower bound for the variances of the estimates obtained from such designs, and that the lower bound is attained for all the estimates, grand mean and main effects, for one particular expansible series. It has also been demonstrated that a much wider class of one-at-a-time designs has the property that the variances of the main-effect estimates attain the lower bound. An upper bound on the efficacy of such designs is $2/(n+1)$.

PERMUTATION-INVARIANT DESIGNS

Properties of the Designs

A class of designs which often provides convenient expansible series is the class of permutation-invariant designs defined in another paper, which is included as Appendix B. For 2^n experiments such a series is characterized by a set of k integers $\{m_1, \cdots, m_k\}$, where for convenience the m_i will be arranged in non-decreasing order. For $n \geq m_k$ factors the design contains those $\binom{n}{m_1}$ treatment combinations with exactly m_1 factors at their high level, plus those $\binom{n}{m_2}$ combinations with exactly m_2 factors at their high level, $\cdots$, plus all those $\binom{n}{m_k}$ combinations with exactly m_k factors at their high level.

Such designs have the property of treating all factors alike. That is, the variances of all main effects are equal, the variances of all two-factor interactions are equal, and covariances between analogous pairs of estimates are equal. The cross-product matrices have a particular pattern which allows general expressions for the elements of the inverse to be obtained. Consider first designs for estimating only the grand mean and main effects, for which the cross-product matrix is as follows:

N	A	A	A	A	A
A	N	B	B	B	B
A	B	N	B	B	B
A	B	B	N	B	B
A	B	B	B	N	B
A	B	B	B	B	N

Although this is only a five-factor example, it serves to illustrate the general case. The three constants N, A, and B are respectively the number of runs, the inner product of a coefficient vector for a main effect with one for the mean, and the inner product of coefficient vectors for two main effects.

In designs for estimating the grand mean, main effects, and two-factor interactions, the cross-product matrices have a somewhat more complex structure, which can be illustrated by a four-factor example:

N	A	A	A	A	B	B	B	B	B	B
A	N	B	B	B	A	A	A	C	C	C
A	B	N	B	B	A	C	C	A	A	C
A	B	B	N	B	C	A	C	A	C	A
A	B	B	B	N	C	C	A	C	A	A
B	A	A	C	C	N	B	B	B	B	D
B	A	C	A	C	B	N	B	B	D	B
B	A	C	C	A	B	B	N	D	B	B
B	C	A	A	C	B	B	D	N	B	B
B	C	A	C	A	B	D	B	B	N	B
B	C	C	A	A	D	B	B	B	B	N

Assigning arabic numbers to the factors and representing interactions by a pair of numbers, the rows and columns of this matrix are ordered so that they are associated with the mean, factors 1 through 4, and interactions in the order 12, 13, 14, 23, 24, 34. Here N, A, and B are defined as before in the part of the matrix

associated with the mean and main effects (the first five rows and columns). Due to the structure of 2^n designs, they also appear elsewhere in the matrix. A appears as the inner product of coefficient vectors for the main effect of a factor and an interaction containing that factor (for example, 1 with 12, 3 with 23, etc.). B is the inner product of coefficient vectors for an interaction and for the grand mean, and of two coefficient vectors for interactions having one factor in common (e.g., 12 with 13), as well. The constant C is defined as the inner product of coefficient vectors for a main effect of one factor and for an interaction involving second and third factors (e.g., 2 with 13). Finally, D is the inner product of coefficient vectors for two interactions involving a total of four factors (13 with 24, etc.).

General expressions for the five constants are as follows for the m-factor design of the general series $\{m_1, \cdots, m_k\}$:

$$N = \Sigma_{i=1}^{k} \binom{n}{m_i}$$

$$A = \Sigma_{i=1}^{k} \left\{\binom{n-1}{m_i-1} - \binom{n-1}{m_i}\right\}$$

$$B = \Sigma_{i=1}^{k} \left\{\binom{n-2}{m_i-2} - 2\binom{n-2}{m_i-1} + \binom{n-2}{m_i}\right\}$$

$$C = \Sigma_{i=1}^{k} \left\{\binom{n-3}{m_i-3} - 3\binom{n-3}{m_i-2} + 3\binom{n-3}{m_i-1} - \binom{n-3}{m_i}\right\}$$

$$D = \Sigma_{i=1}^{k} \left\{\binom{n-4}{m_i-4} - 4\binom{n-4}{m_i-3} + 6\binom{n-4}{m_i-2} - 4\binom{n-4}{m_i-1} + \binom{n-4}{m_i}\right\} ,$$

with the convention that when a binomial coefficient $\binom{p}{q}$ is undefined because $q < 0$ or $p < q$, its value is zero. The derivation for B will be given, the others being completely analogous.

From an examination of the coefficient matrix X, given any two factors, one can see that B is the number of occurrences of 00 minus twice the number of occurrences of 01 plus the number of occurrences of 11. Considering only those treatment combinations with m_i factors at their high levels, there are exactly $\binom{n-2}{m_i-2}$ combinations with two factors fixed at levels 1 and 1, because there are m_i-2 more 1's to be assigned to the n-2 remaining factors. Similarly, the number of occurrences of 0 and 1 for two fixed factors is $\binom{n-2}{m_i-1}$, and the number of occurrences of 0 and 0 is $\binom{n-2}{m_i}$. Combining these and summing over i gives the above expression.

Covariance Matrices of Permutation-Invariant Designs

The covariance matrix (the inverse of the cross-product matrix) for permutation-invariant designs also exhibits a distinctive pattern. For main-effect designs the inverse has the following form, again illustrated by a five-factor example:

α	β	β	β	β	β
β	γ	δ	δ	δ	δ
β	δ	γ	δ	δ	δ
β	δ	δ	γ	δ	δ
β	δ	δ	δ	γ	δ
β	δ	δ	δ	δ	γ

The constant α times σ^2 (the constant error-variance of the errors in the observations) is the variance of the estimate of the mean, $\gamma\sigma^2$ is the variance of an estimated main effect, $\beta\sigma^2$ is the covariance between the estimated mean and an estimated effect, and $\delta\sigma^2$ is the covariance between the estimates of two main effects.

Because of the patterned structure of these matrices, a set of five distinct equations in the four unknowns α, β, γ, and δ can be obtained by formally multiplying the cross-product matrix by the covariance matrix. For the general case of n factors the equations are as follows:

$$
\begin{aligned}
&N\alpha + n A\beta = 1 \\
&A\alpha + [N + (n-1)B]\beta = 0 \\
&N\beta + A\gamma + (n-1)A\delta = 0 \\
&A\beta + N\gamma + (n-1)B\delta = 1 \\
&A\beta + B\gamma + [N + (n-2)B]\delta = 0 \quad .
\end{aligned}
$$

These equations are consistent, and the explicit solutions are as follows:

$$
\begin{aligned}
\alpha &= [N + (n-1)B]/\Delta \\
\beta &= -A/\Delta \\
\gamma &= [(n-1)(NB-A^2) + N(N-B)]/(N-B)\Delta \\
\delta &= -(NB-A^2)/(N-B)\Delta \quad ,
\end{aligned}
$$

where Δ is given by $\Delta = n(NB-A^2) + N(N-B)$.

The covariance matrix for the estimation of the mean, main effects, and two-factor interactions is patterned like the cross-product matrix for this case. Its elements will be denoted by small Latin letters. The four-factor example is as follows:

a	b	b	b	b	c	c	c	c	c	c
b	d	e	e	e	f	f	f	g	g	g
b	e	d	e	e	f	g	g	f	f	g
b	e	e	d	e	g	f	g	f	g	f
b	e	e	e	d	g	g	f	g	f	f
c	f	f	g	g	h	i	i	i	i	j
c	f	g	f	g	i	h	i	i	j	i
c	f	g	g	f	i	i	h	j	i	i
c	g	f	f	g	i	i	j	h	i	i
c	g	f	g	f	i	j	i	i	h	i
c	g	g	f	f	j	i	i	i	i	h

Again the matrix and its inverse may be formally multiplied to obtain a set of fourteen consistent equations in the ten unknowns. These equations are as follows:

$$Na + nAb + \binom{n}{2}Bc = 1$$

$$Aa + [N+(n-1)B]b + [(n-1)A + \binom{n-1}{2}C]c = 0$$

$$Ba + [2A+(n-2)C]b + [N+2(n-2)B + \binom{n-2}{2}D]c = 0$$

$$Nb + Ad + (n-1)Ae + (n-1)Bf + \binom{n-1}{2}Bg = 0$$

$$Ab + Nd + (n-1)Be + (n-1)Af + \binom{n-1}{2}Cg = 1$$

$$Ab + Bd + [N+(n-2)B]e + [A+(n-2)C]f + [(n-2)A + \binom{n-2}{2}C]g = 0$$

$$Bb + Ad + [A+(n-2)C]e + [N+(n-2)B]f + [(n-2)B + \binom{n-2}{2}D]g = 0$$

$$Bb + Cd + [2A+(n-3)C]e + [2B+(n-3)D]f + [N+2(n-3)B + \binom{n-3}{2}D]g = 0$$

$$Nc + 2Af + (n-2)Ag + Bh + 2(n-2)Bi + \binom{n-2}{2}Bj = 0$$

$$Ac + (N+B)f + (n-2)Bg + Ah + (n-2)(A+C)i + \binom{n-2}{2}Cj = 0$$

$$Ac + 2Bf + [N+(n-3)B]g + Ch + [2A+2(n-3)C]i + [(n-3)A+\binom{n-3}{2}C]j = 0$$

$$Bc + 2Af + (n-2)Cg + Nh + 2(n-2)Bi + \binom{n-2}{2}Dj = 1$$

$$Bc + (A+C)f + [A+(n-3)C]g + Bh + [N+(n-2)B+(n-3)D]i + [(n-3)B+\binom{n-3}{2}D]j = 0$$

$$Bc + 2Cf + [2A+(n-4)C]g + Dh + [4B+2(n-4)D]i + [N+2(n-4)B+\binom{n-4}{2}D]j = 0$$

Attempts have been made to obtain explicit solutions, but the algebraic complexity precludes obtaining simple expressions for the inverse elements. It is far easier to evaluate and solve the above equations for a particular case, and a simple computer routine has been written to do this.

NUMERICAL RESULTS ON RESOLUTION-FIVE DESIGNS

Numerical results have been obtained for a large number of resolution-five permutation-invariant designs, and these are given in Table I. The table lists the series (that is, the m_i's); number of factors; the constants M, A, B, C, and D; the efficiency of the estimates of the mean, main effects, and interactions; the prediction index; the degree-of-freedom efficiency; and the efficacy. An examination of the table reveals that the best series in terms of efficacies of those tried are as follows for various values of the number of runs guaranteed to be made.

guaranteed minimum	best series	n	efficacy in %
4	{0,1,2}	2	100
		3	57
		4	23
		5	10
7	{1,2,3}	3	57
		4	52
		5	29
		6	14
11	{1,2,4}	4	67
		5	69
		6	50
		7	28
16	{1,3,5}	5	100
		6	69
		7	45
22	{0,2,5}	6	89
		7	58
		8	35

TABLE I

EVALUATION OF PERMUTATION-INVARIANT DESIGNS

Series	n	N	A	B	C	D	Efficiency of mean	Efficiency of main effects	Efficiency of interactions	P.I.	d.f. efficiency	efficacy
{0,1,2}	2	4	0	0	0	0	100	100	100	100	100	100
	3	7	-1	-1	-1	-1	57	57	57	57.14	100	57
	4	11	-3	-1	1	3	10	18	36	22.86	100	23
	5	16	-6	0	2	0	2	6	25	9.63	100	10
	6	22	-10	2	2	-2	1	3	18	4.62	100	5
{0,1,3}	5	16	-2	0	-2	0	64	45	64	56.69	100	57
	6	27	-5	-1	-1	3	9	15	45	25.56	81.48	21
	7	43	-11	-1	1	3	2	5	31	10.82	67.44	7
{0,2,3}	4	11	1	-1	-3	3	94	65	65	67.29	100	67
	5	21	-1	-3	-1	5	49	57	58	57.39	76.19	44
	6	36	-6	-4	2	4	13	27	46	35.60	61.11	22
	7	67	-15	-3	5	1	38	45	69	59.32	43.28	26
{1,2,3}	3	7	1	-1	1	-1	57	57	57	57.14	100	57
	4	14	0	-2	0	-2	57	91	57	66.17	78.57	52
	5	25	-3	-3	1	1	23	49	49	45.91	64.00	29
	6	41	-9	-3	3	1	6	18	40	25.25	53.66	14
	7	63	-19	-1	5	-1	2	7	33	14.82	46.03	7
{0,1,4}	6	22	0	2	-4	-2	55	54	65	61.54	100	62
	7	43	-1	-1	-5	3	62	19	52	54.69	67.44	37
	8	79	-7	-5	-3	7	4	6	34	11.29	46.84	5
{0,2,4}	5	16	0	0	0	0	100	100	100	100	100	100
	6	31	-1	-1	-1	-1	94	94	94	93.84	70.97	67
	7	57	-5	-3	-1	1	47	62	80	73.35	50.88	37
	8	85	-9	-5	1	3	57	75	84	80.97	43.53	35
{0,3,4}	5	16	4	0	-4	0	64	28	45	38.50	100	38
	6	36	4	-4	-4	4	74	54	54	54.72	61.11	33
	7	71	-1	-9	-1	7	38	56	51	51.33	40.85	21
	8	127	-15	-13	5	7	12	30	44	48.39	29.13	14
{1,2,4}	4	11	-1	-1	3	3	94	65	65	67.29	100	67
	5	20	-2	0	2	-4	95	91	83	86.23	80.00	69
	6	36	-4	0	0	-4	88	86	79	81.05	61.11	50
	7	63	-9	-1	-1	-1	37	47	71	61.51	46.03	28
	8	106	-20	-2	0	2	8	20	59	37.04	34.91	13
{1,3,4}	5	20	2	0	-2	-4	95	91	83	86.23	80.00	69
	6	41	1	-3	-3	1	81	77	83	80.98	53.66	43
	7	77	-5	-7	-1	5	33	52	70	62.27	37.66	23
	8	134	-20	-10	4	6	10	28	57	42.22	27.61	12

TABLE I (Continued)

Series	n	N	A	B	C	D	Efficiency of mean	Efficiency of main effects	Efficiency of interactions	P.I.	d.f. efficiency	efficacy
{2,3,4}	4	11	3	-1	-1	3	10	18	36	22.86	100	23
	5	25	3	-3	-1	1	23	49	49	45.91	64.00	29
	6	50	0	-6	0	2	23	86	49	52.63	44.00	23
	7	91	-9	-9	3	3	10	45	45	42.22	32.87	14
	8	154	-28	-10	8	2	4	17	40	26.19	24.03	6
{0,1,5}	7	29	3	5	-5	-3	12	52	57	49.40	100	49
	8	65	7	1	-9	1	19	19	52	36.91	56.92	21
	9	136	6	-8	-10	8	45	6	32	17.16	33.82	6
{0,2,5}	6	22	-2	2	1	-2	86	92	88	89.21	100	89
	7	43	-1	3	-1	-5	80	97	84	86.57	67.44	58
	8	85	-1	1	-5	-3	99	74	83	81.30	43.52	35
{0,3,5}	6	27	3	-1	-1	3	88	86	82	83.57	81.48	68
	7	57	3	-3	-1	1	88	95	95	94.38	50.87	48
{0,4,5}	6	22	8	2	-4	-2	35	11	31	20.47	100	20
	7	57	13	-3	-7	1	47	26	45	38.45	50.87	20
{1,2,5}	5	16	-4	0	4	0	64	28	45	38.50	100	38
	6	27	-5	3	3	-5	58	74	72	72.11	81.48	59
	7	49	-5	5	-1	-7	61	91	74	76.76	59.18	45
	8	92	-6	4	-6	-4	91	65	78	75.44	40.21	30
{1,3,5}	5	16	0	0	0	0	100	100	100	100	100	100
	6	32	0	0	0	0	100	100	100	100	68.75	69
	7	63	-1	-1	-1	-1	99	99	99	98.77	46.03	45
{1,4,5}	6	27	5	3	-3	-5	59	74	72	72.11	81.48	59
	7	63	9	-1	-7	-1	81	65	85	79.09	46.03	36
{2,3,5}	5	21	1	-3	1	5	49	57	58	57.29	76.19	44
	6	41	-1	-3	3	1	81	77	83	80.92	53.65	43
	7	77	-5	-3	3	-3	85	93	84	85.78	37.66	32
{2,4,5}	5	16	2	0	2	0	64	45	64	56.68	100	57
	6	36	4	0	0	-4	88	86	79	81.05	61.11	50
	7	77	5	-3	-3	-3	85	93	84	85.78	37.66	38
{0,0,2,3}	4	12	0	0	-4	4	100	62	62	63.78	91.66	58
{0,1,2,4}	5	21	-3	1	1	-3	90	90	90	90.30	76.19	69
{0,1,3,4}	5	21	1	1	-3	-3	94	88	88	91.39	76.19	70
{0,0,3,5}	6	28	2	0	-2	4	97	87	80	82.81	78.57	65
	7	58	2	-2	-2	2	97	96	96	95.87	50.00	48
{0,1,4,5}	6	28	4	4	-4	-4	57	72	72	70.97	78.57	56
{0,2,3,5}	6	42	-2	-2	2	2	95	94	94	94.22	52.38	49
{0,2,4,5}	5	17	1	1	1	1	97	97	97	97.15	94.11	91
	6	37	3	1	-1	-3	95	96	95	95.18	59.45	57

The first of these resolution-5 series, $\{0,1,2\}$, is a direct analogue of the class of one-at-a-time designs for the resolution-3 case. The series is fully contractible, and each design contains only as many runs as there are parameters to be estimated. Like one-at-a-time designs, the series becomes inefficient very rapidly as the number of factors increases.

For any given series, the efficiencies tend to drop monotonically as the number of factors increases, as one might expect. It is noteworthy, however, that the efficiency of the mean consistently drops much more rapidly than the efficiency of the main effects, which in turn drops more rapidly than the efficiency of the interactions. Thus, if interactions are of primary interest, the series are particularly attractive.

Some interesting irregularities in the monotonicity of the efficiencies as the number of factors increases may be noted, however. For example, in the series $\{0,2,4\}$ the efficiencies of the mean, main effects, and interactions all decrease regularly as n goes from 5 to 6 to 7, and then jump up as n increases from 7 to 8; similarly they jump for the series $\{0,2,3\}$ as n goes from 6 to 7. There seems to be no consistent pattern for these irregularities, but rather they appear to be an artifact of the particular numerical values of the constants associated with the designs. For some of the later series in the table, for which $m_k = 5$, there is a tendency for the efficiencies as measured by the prediction index to increase and then peak. When this is observed the degree-of-freedom efficiency is very low at the peak, so that these series do not appear to be useful.

For some designs it appears from the constants A through D that the efficiencies can be improved by appending a run at the treatment combination with all factors at their low level. This procedure adds 1 to M, B, and D and subtracts

1 from A and C. The last designs in the table were constructed in this way, and for most of these selected designs the prediction indices did in fact increase to some degree. The value of the extra run, as measured by the efficacy, was only slight, however, except for one design. The $\{2,4,5\}$ series with $n = 5$ has an efficacy of 57%, while the series $\{0,2,4,5\}$ has an amazingly higher efficacy of 91%.

THE DERIVATION OF EXPERIMENTAL DESIGNS BY LINEAR PROGRAMMING

GENERAL APPROACH

Various classes of incomplete factorial designs can be characterized as solutions to sets of linear constraints. Papers have been written giving the characterizations for orthogonal designs (Appendix A), and for the permutation-invariant and clumpwise-orthogonal classes of non-orthogonal designs (Appendix B). In the case of orthogonal designs, the characterization is somewhat different from that given by Addelman [1]. An investigation of the differences between the two characterizations has revealed that designs enjoy two optimality properties if they are orthogonal in the sense given in Appendix A; these results are described in Appendix C.

Given the number of factors and the interactions to be estimated, the linear constraints characterizing the class of designs can be used to find the particular design in the class which has the smallest number of runs. It is always possible to find a design in the class of interest, but those designs which are readily obtainable are often too large to be practical. Starting from the known large design, linear programming can be applied in order to obtain the smallest design of the class. This procedure is illustrated in Appendix A for the family of orthogonal 2^4 resolution 3 designs, where the half-replicate is obtained starting from the full factorial.

In order to study the linear-programming technique for deriving designs, a special-purpose linear-programming computer routine has been written. The novel features of the routine, which were based on experience with particular experimental-design problems, are discussed below. Unlike the product-form algorithms

in general use for solving linear programming problems, neither the basis matrix nor its inverse, but only the triangularized basis, is available at any iteration. Although the routine has been used only for experimental-design problems, it can presumably be used to advantage for any problem of a combinatorial nature such as assignment problems, transportation problems, and scheduling problems.

DESCRIPTION OF THE LINEAR PROGRAMMING COMPUTER ROUTINE

The Simplex Method

A brief review of the simplex method will be useful in order to introduce notation. The constraint matrix will be denoted by A, and its number of rows and columns by N and M, respectively, where $N < M$. The constraints are represented by the matrix equation $Aw = b$, where w is the M-component vector of variables and b is the N-component vector of right-hand sides. For convenience and without loss of generality the linear form to be minimized will be taken to be simply the value of the first component of w, which will be denoted by λ. Any basis selected from the columns of A will be designated by B. The corresponding basic solution w can be constructed from $B^{-1}b = p$ as follows: if column A_j appears in the basis as column B_i, then $w_j = p_i$; if column A_j does not appear in the basis, then $w_j = 0$.

Suppose the solution corresponding to a particular basis B is feasible. Let $B_1 = A_1$, so that $p_1 = w_1$, and let $G = B^{-1}A$. Any column of A can be expressed as a linear combination of columns of B in the form $A_j = g_{1j}B_1 + g_{2j}B_2 + \cdots + g_{Nj}B_N$. Suppose we wish to change the basis by introducing say A_j and deleting B_r. Expressing B_r in terms of the new basis, we have

$$B_r = -\frac{g_{1j}}{g_{rj}} B_1 \cdots\cdots \frac{g_{(r-1)j}}{g_{rj}} B_{r-1} - \frac{g_{(r+1)j}}{g_{rj}} B_{r+1} \cdots\cdots \frac{g_{Nj}}{g_{rj}} B_N + \frac{1}{g_{rj}} A_j ,$$

assuming $g_{rj} \neq 0$. If we substitute this value for B_r in the equation $\sum_{i=1}^{N} p_i B_i = b$, we have $\sum_{i \neq r}(p_i - p_r \frac{g_{ij}}{g_{rj}})B_i + \frac{p_r}{g_{rj}} A_j = b$. From this equation can be determined immediately the following elements of the solution p^* corresponding to the new basis: $p_i^* = p_i - p_r \frac{g_{ij}}{g_{rj}}$, for $i \neq r$, and $p_r^* = p_r/g_{rj}$.

Suppose we wish to make the choice of a column A_j to enter the basis in such a way that λ does not increase. Since λ is the value of the first element of the solution, it changes from p_1 to $p_1^* = p_1 - g_{1j}p_r/g_{rj}$, so that it will not increase if g_{1j} and g_{rj} have the same sign. Assuming that $p_i \geq 0$ so that the original solution is feasible, suppose we wish to choose B_r in such a way that feasibility is maintained. In order for p_r^* to be positive g_{rj} must be positive, and in order for the remaining p_i^* to be positive, r must be selected so that $p_r/g_{rj} = \theta$ is the minimum value of those p_i/g_{ij} for which g_{ij} is positive.

Thus, if j is chosen so that g_{1j} is positive, and if r is chosen so that θ is minimized, then the solutions associated with successive bases remain feasible and the value of λ cannot increase. Thus, the process may be continued iteratively until finally no column can be found to enter the basis.

Properties of Experimental Design Problems

The linear programming problems for finding incomplete factorial designs have several distinctive characteristics. These are the large size of the constraint matrices, the fact that the matrices consist primarily of 0's and -1's (see Appendix A), and the fact that the starting basic feasible solutions are often highly degenerate. Due to the degeneracy, the course of the simplex method is typically as follows: many iterations are performed with no change in λ because a p_i is zero for which g_{ij} is positive, and thus $\theta = 0$; then on successive iterations the value λ decreases gradually, with the associated solutions being

nondegenerate; finally a final degenerate solution is obtained, but many more iterations are performed before the condition that no g_{1j} is positive occurs.

Triangularization in the Routine

Execution is initiated with the matrix A stored in core and on auxiliary tape, with the first N columns of A being the basis. The basis matrix is triangularized by Gaussian elimination with the elimination calculations applied to the remaining columns of A and to b, the vector of right-hand sides. It will be helpful in the sequel to go through the elimination steps for a simple example. Suppose the original matrix is denoted by

$$\begin{matrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{matrix} ,$$

where the fourth column will represent a column of A not in the basis and the fifth column will represent the right-hand sides. After one step the matrix becomes

$$\begin{matrix} b_{11} & b_{12} & 0 & b_{14} & b_{15} \\ b_{21} & b_{22} & 0 & b_{24} & b_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{matrix} ,$$

where $b_{ij} = a_{ij} - a_{i3}a_{3j}/a_{33}$, assuming that division by a_{33} is possible. The next step yields the triangularized system

$$\begin{matrix} c_{11} & 0 & 0 & c_{14} & c_{15} \\ b_{21} & b_{22} & 0 & b_{24} & b_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{matrix} ,$$

where $c_{ij} = b_{ij} - b_{i2}b_{2j}/b_{22}$, again assuming that division by b_{22} is possible. In the program the divisors, or pivots, are checked to see that they are larger than a tolerance t_1 (usually 10^{-2} or 10^{-3}). If not, rows are interchanged until a large enough pivot is found. If none can be found, the matrix is considered singular. Because the matrices have as elements 0's and -1's, they are relatively well conditioned. The program was written under the assumption that the process of picking any pivot larger than the tolerance, rather than a time-consuming search for the largest pivot, would be satisfactory. Some difficulties due to roundoff have been encountered (see below), but they may be due to the size of the basis matrices rather than an inherent weakness in this method.

Rules for Selecting a Column to Enter the Basis

After triangularization, the selection of a column to enter the basis is a relatively simple matter, since the elements g_{1j} are just c_{1j}/c_{11}. Ordinarily there will be a number of columns for which g_{1j} is positive, and three rules have been tried for deciding among them. These are:

Rule 1: Select the first column (starting with $j = N+1$) for which g_{1j} is positive.

Rule 2: Select the column with the largest positive g_{1j}.

Rule 3: Select a column for which $\theta > 0$ if one exists, otherwise use Rule 2.

An experimental investigation was performed in order to evaluate these rules. Two representative problems, deriving the smallest 2^6 and 2^7 resolution 3 designs, were used for this study. The constraint matrices for these problems were of size 23x66 and 30x130, respectively. For each problem ten decks were constructed, with the columns of A arranged in a different random order in each deck, and all three rules were used with each deck. The number of iterations before the final solution was obtained and number of seconds of execution time were as follows:

RUNNING TIME AND NUMBER OF ITERATIONS FOR THREE SELECTION RULES

Problem	2^6						2^7					
	Rule 1		Rule 2		Rule 3		Rule 1		Rule 2		Rule 3	
Deck	sec.	iter.	sec.	iter.	sec.	iter.	sec.	iter.	sec.	iter.	sec.	iter.
1	39	50	9	26	14	49	269	1075	41	66	82	164
2	15	52	8	25	10	28	53	137	56	96	80	190
3	14	69	28	81	17	61	74	217	69	165	136	311
4	11	48	10	35	12	50	198	727	117	351	159	319
5	10	39	9	25	11	36	69	217	290	587	99	176
6	14	71	9	32	16	80	77	307	44	77	86	168
7	14	68	9	33	12	43	122	399	88	176	267	752
8	24	161	8	28	11	36	160	605	351*	677	125	238
9	11	59	18	26	12	44	325*	1630	172	401	44	66
10	10	49	17	29	29	113	81	464	47	60	155	276
Averages	16.2	66.6	12.5	34.0	14.4	54.0	142.8	577.8	127.5	265.6	121.3	266.0

The number of seconds of execution is the important variable, and these data have been analyzed. As a check, three of the runs were duplicated, and the running time was exactly the same twice and was one second different once. For two of the 2^7 runs, labeled with an asterisk in the above table, execution was stopped before a solution had been reached, and the time at which this occurred was entered in the table. The distribution of times is highly skewed, so an analysis of variance was performed on the double logarithm of the times. This analysis revealed that neither the decks nor the rules were significantly different when compared with the rule by deck interaction. Therefore, for these two problems there is no clear basis for choosing between the three rules. For a particular deck there is often a considerable difference between the three rules, however, but this fact is of no practical use since for an actual problem the properties of the particular arrangement of the columns are unknown. For other problems than the 2^6 or 2^7 it may sometimes be possible to make an intuitive choice based on the nature of the rules. For example, if the number of columns not in the basis is small compared with the number which are in, the amount of additional computation required for applying Rule 2 is comparatively small, and one might wish to apply that rule.

Rules for Selecting a Column to Leave the Basis

Once a column, say the j-th, has been selected to enter the basis, a column must be selected to leave. Due to degeneracy, the value of θ is usually 0, and moreover there are typically several columns which have this value of θ. There is nothing lost, and much computational time saved, by selecting the column to leave with the lowest order number. If column r yields the value $\theta = 0$, then the p_i and g_{ij} only need be calculated for $i = 1, \cdots, r$. Moreover, when the new column A_j is inserted into the basis, the basis is still triangularized except for the first r rows and columns, which produces a substantial saving in computer time over triangularizing the whole basis. Since it is highly advantageous to pick low r's, the matrix A should be initially arranged so that columns of B corresponding to zero elements of p are placed in the basis immediately after A_1, the column corresponding to λ.

The column which has been selected to enter the basis is substituted for the column to leave, and the latter is set equal to zero. In the course of applying the simplex method it may be necessary for a column which has left the basis to reenter at a later iteration. Because of this, in most linear programming routines a column which has left the basis is immediately updated in such a way that it enjoys the same computational status as if it had never been in the basis. Because of the size of the constraint matrices, there is usually a wealth of columns which can enter the basis at any iteration, so that little is lost by deleting the columns which have been removed. When it is determined that no column can enter the basis, the original matrix is called back in from auxiliary tape storage, so that the columns which had been deleted can become available for reentry.

Detection of Roundoff Error

During the course of repeated retriangularizations of the basis matrix, roundoff error inevitably builds up. If this buildup is ignored, singular bases or infeasible solutions may be obtained, and much computer time lost in meaningless calculations. Suppose that a system of equations has been triangularized, and that the solutions to the system are to be obtained by back substitution. An example of the triangularized system was shown above to be of the form

$$\begin{matrix} c_{11} & 0 & 0 & c_{14} & c_{15} \\ b_{21} & b_{22} & 0 & b_{24} & b_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{matrix} \quad ,$$

where the fifth column of the original matrix is the vector of right-hand sides. We wish to solve for $p = (p_1, p_2, p_3)'$ such that

$$A_1 p_1 + A_2 p_2 + A_3 p_3 = A_5 \quad ,$$

where the A_j are the original columns of the matrix. From the triangularized system we have immediately

$$\begin{aligned} p_1 &= c_{15}/c_{11} \\ p_2 &= (b_{25} - b_{21}p_1)/b_{22} \\ p_3 &= (a_{35} - a_{31}p_1 - a_{32}p_2)/a_{33} \quad . \end{aligned}$$

Note that the computation of p_3 involves all the elements which have previously been computed, so that this element is the most sensitive to the accumulation of roundoff error. Once again the degeneracy of the solutions, usually considered a disadvantage in linear programming, can be put to good use. Due to degeneracy, many of the p_i have the value zero except for roundoff error, which can be

expected to make these p_i small positive or small negative numbers. Since it is not possible for a p_i to have a negative value legitimately, a test is made of all the p_i which are calculated at a given iteration. If any is found which is a larger negative number than a specified tolerance t_2 (usually from -10^{-3} to -10^{-5}), the solution is considered to be infeasible, even though the infeasibility may be due solely to roundoff error.

If a solution is encountered which is considered to be infeasible, or if a basis matrix is found which is considered singular, an attempt is made to correct the problem automatically. This is done by going back one iteration, calling the constraint matrix back from tape storage, and performing a new complete triangularization on it. On many occasions this procedure has successfully corrected the difficulty. If the matrix is found to be singular or the solution infeasible after going back one iteration and starting with original data, the problem is discontinued.

COMPUTATIONAL EXPERIENCE WITH THE ROUTINE

Besides the 2^6 and 2^7 problems mentioned earlier, attempts to determine explicit 2^8, 2^9, and 3^5 designs have been attempted without complete success. For the 2^8 and 3^5 the value of λ decreased to the minimum theoretical values of 3 and 2, respectively, but the solution did not consist of integers. This contradicts previous experience noted in Appendix A; previously no non-integral solution in which λ was an integer had been obtained for an experimental-design problem.

In the class of orthogonal 2^8 resolution-3 designs the smallest is known to be the twelve-run Plackett-Burman design [2]. Attempts were made to derive this design starting from the twenty-run Plackett-Burman design, for which $\lambda = 5$. The constraint matrix for this problem is of size 38×258. The basis matrix was

called back from tape and retriangularized for four different reasons during the course of this problem; these were(a)an observed singular basis matrix,(b)an observed infeasibility,(c)no column could enter the basis, or(d)100 iterations were performed since the last callback. The complete history of this problem, involving three separate computer runs, is as follows.

Callback	Iteration	Comment
0	24	singularity
1	100	
2	94	no column can enter
3	100	
4	45	false infeasibility (due to roundoff)
5	100	
6	51	true infeasibility
7	1	still infeasible, $\lambda = 6.247$
7	3	initial solution reobtained ($\lambda = 5$)
7	94	false infeasibility
8	100	
9	56	true infeasibility
10	1	still infeasible, $\lambda = 12.000$
10	2	initial solution reobtained
10	79	singularity
11	68	false infeasibility
12	7	first decrease in λ to $\lambda = 4.578$, followed by 10 consecutive changes
12	17	false infeasibility, $\lambda = 3.540$
13	1	$\lambda = 3.540$, followed by 38 consecutive changes
13	39	false infeasibility, $\lambda = 2.881$
14	1	$\lambda = 2.876$, followed by 26 consecutive changes
14	27	false infeasibility, $\lambda = 2.576$
15	0	singularity or infeasibility after callback
		NEW RUN
0	2	$\lambda = 2.580$, followed by 6 consecutive changes
0	8	no column can enter, $\lambda = 2.517$
1	0	minimum solution reached, $\lambda = 2.500$
		NEW RUN, start at $\lambda = 3.027$, require $\lambda \geq 3.000$
0	1	$\lambda = 3.000$, no column can enter
1	0	minimum solution reached, $\lambda = 3.000$

The variables w_i in the linear-programming formulation of these problems correspond to the treatment combinations of the full factorial. The numerical value of a w_i is the number of times the corresponding treatment combination is

included in the design. The following final solution, obtained after about 1000 iterations, does not consist of integers, hence is useless as a design.

Treatment Combination	Value	Treatment Combination	Value
0 0 0 0 0 0 0 0 0	1.000	1 0 0 1 0 0 1 0	.453
0 0 0 1 1 1 1 1 0	.720	1 0 0 1 1 1 1 1 0	.087
0 0 1 0 1 0 0 0 1	.466	1 0 1 0 0 0 0 0 1	.098
0 0 1 0 1 1 0 1	.058	1 0 1 0 0 1 0 0	.125
0 0 1 1 0 1 1 0	.172	1 0 1 0 0 1 0 1	.108
0 0 1 1 1 1 1 0 1	.548	1 0 1 0 0 1 1 0	.060
0 0 1 1 1 1 1 1 0	.035	1 0 1 0 1 1 0 0	.346
0 1 0 0 0 0 0 1 1	.407	1 0 1 1 0 0 0 1	.047
0 1 0 0 0 1 1 0	.008	1 0 1 1 0 0 1 1	.936
0 1 0 0 1 0 1 0	.035	1 1 0 0 0 0 0 0 0	.058
0 1 0 0 1 0 1 1	.381	1 1 0 0 1 1 1 0	.323
0 1 0 1 0 1 0 1	.449	1 1 0 0 1 1 1 1	.050
0 1 1 0 0 1 1 1	.646	1 1 0 1 0 1 0 0	.315
0 1 1 1 0 1 1 0	.318	1 1 0 1 0 1 0 1	.309
0 1 1 1 1 0 0 0	.479	1 1 0 1 1 0 0 1	.666
0 1 1 1 1 0 1 0	.233	1 1 1 0 0 1 0 0	.490
0 1 1 1 1 1 1 1	.045	1 1 1 0 1 0 1 0	.555
1 0 0 0 1 1 0 1	.203	1 1 1 0 1 1 0 1	.466
1 0 0 0 1 1 1 1	.536	1 1 1 1 1 0 0 0	.186

The smallest orthogonal 2^9 resolution-3 design has not as yet been obtained, although three computer runs have been attempted. The constraint matrix is of size 47×514, and the computations were initialized with the 24-run Plackett-Burman design, for which $\lambda = 6$. The first computer run went only 33 iterations with two callbacks before stopping. The second went only 10. On the third run, the value of λ decreased from 6 to 4.709 on the 74th iteration. Another 9 iterations brought the value down to $\lambda = 3.460$. In this case roundoff error during the new triangularization built up to such a degree that false infeasibility was observed immediately and the problem was discontinued. No further runs have been attempted with the 2^9.

The smallest known orthogonal 3^5 design of resolution 4 is the 1/3 replicate involving 81 runs. It follows from the results of Appendix A that the number of runs in such a design must be a multiple of 27, and it is known that 27 runs is insufficient. It appears, therefore, that there should exist a 54-run design. The constraint matrix is of size 132 x 245, but due to its size, the portion of the matrix in core at any one time is 132 x 189, the remainder being on auxiliary tape storage. Six computer runs were performed, involving 381 iterations and 27 callbacks, each of which stopped when singularity or infeasibility was observed immediately after a callback. In between runs columns were rearranged or the tolerances t_1 and t_2 defined above were changed in order that the problem would not halt as it had during the previous run. During the seventh computer run, involving 120 iterations and 6 callbacks, the value of λ decreased to 2.075, and during the eighth run the value $\lambda = 2.000$ was obtained for the following solution:

Solution to 3^5 Problem--Treatment Combinations and the Corresponding Values of the Solution

00000	1.000	01112	.247	02210	.409	10221	.270	12100	.521	20211	.330
00012	.318	01122	.469	02211	.148	10222	.312	12102	.341	20220	.734
00021	.637	01200	.443	02220	.317	11001	.451	12111	.583	21000	.734
00022	.045	01201	.283	02221	.334	11002	.419	12120	.231	21010	.303
00102	.683	01202	.167	10001	.413	11010	.461	12122	.323	21012	.230
00110	.016	01210	.258	10002	.055	11011	.084	12201	.475	21021	.734
00111	.482	01220	.146	10010	.307	11022	.585	12210	.277	21102	.633
00120	.820	01221	.353	10011	.095	11100	.615	12212	.633	21111	.633
00201	.091	01222	.351	10012	.545	11102	.040	12221	.308	21120	.734
00202	.226	02001	.604	10020	.262	11112	.548	12222	.306	21201	.633
00210	.164	02010	.498	10021	.323	11121	.734	20001	.532	21210	.230
00211	.790	02011	.028	10100	.037	11122	.064	20010	.431	21212	.605
00212	.230	02012	.233	10101	.446	11200	.130	20012	.303	21222	.532
00222	.499	02022	.637	10110	.596	11202	.345	20022	.734	22002	.734
01002	.396	0210[illegible]	.216	10112	.274	11210	.176	20100	.532	22011	.734
01011	.551	02101	.389	10121	.238	11211	.731	20102	.303	22020	.532
01012	.371	02110	.082	10122	.410	11220	.618	20111	.303	22101	.532
01020	.502	02112	.601	10200	.431	12000	.266	20112	.330	22110	.734
01021	.179	02120	.215	10201	.215	12002	.396	20121	.532	22122	.734
01100	.079	02121	.496	10202	.403	12011	.507	20201	.303	22200	.734
01101	.633	02200	.262	10201	.183	12020	.704	20202	.330	22212	.532
01110	.572	02202	.529	10220	.185	12021	.127	20210	.303	22221	.734

Attempts were made to obtain a solution by holding λ fixed and minimizing an arbitrarily selected component of the solution vector, but they did not succeed in obtaining an integral solution. One approach, as yet untried, which may be useful is to minimize the sum of an arbitrary selection of the w_i.

One special run was conducted on the 3^5, during which running times and information on the progress of the routine were printed rather than the usual output. The maximum number of iterations between callbacks was set at 30. A total of 15 callbacks was made, many of which ran the full 30 iterations. The amount of time required for a complete triangularization of the 132×189 system averaged 15 seconds. It took about 3 seconds to call the matrix back from tape. For the first 29 iterations, up till the first callback, the column numbers of the columns which left the basis were as follows:

2	4	21	15	5	9
3	2	4	30	6	4
3	6	23	13	3	10
4	14	6	40	8	2
2	16	2	4	8	

For the largest retriangularization, involving the first 40 rows and columns, the time was 53/60 of one second. The smallest retriangularizations were performed in 1/60 of one second. This illustrates the advantage of allowing the earliest possible column to leave the basis. It should be noted that the amount of time taken for all computations other than retriangularizations and reading in from tape is negligible. The total execution time for this run on the I.B.M. 7090 computer was 6 minutes and 40 seconds.

Although experience with the routine has not been entirely successful, it has demonstrated the general feasibility of the techniques used. In particular, the use of the triangularized matrix rather than the basis and its inverse appears to be advantageous. The method used to eliminate roundoff error works quite well, as is evidenced by the number of times false infeasibilities were detected and corrected. The number of iterations after callback at which false infeasibility is detected varies considerably from callback to callback. Therefore, a constant callback frequency would not appear to be effective. It is interesting to note that more success was had with the 3^5, with constraint matrix of order 132, than with the 2^9, whose constraint matrix is only of order 47, at least in the sense that more iterations per computer run were performed. It is gratifying that a problem as large as the 3^5 could be handled with even partial success, but it is disappointing that the 2^9 turned out to be so difficult a problem.

It is unlikely that linear programming will prove to be a useful technique for deriving new practical incomplete factorial designs due to the sizes of the constraint matrices involved, the number of iterations required for solutions, and the problem of non-integral solutions. It does appear, however, that it may be possible to use the linear constraints for these problems to direct a systematic trial-and-error search for designs on a computer. It is hoped that the special methods incorporated into the computer routine will have more general usefulness with other types of linear programming problems.

REFERENCES

[1] ADDELMAN, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. Technometrics 4, 21-46.

[2] PLACKETT, R. L. and BURMAN, J. P. (1946). The design of optimum multifactorial experiments. Biometrika 33, 305-325.

[3] WEBB, S. R. and GALLEY, S. W. (1965). A computer routine for evaluating incomplete factorial designs. To appear as a Technical Documentary Report, Aerospace Research Laboratories, Wright-Patterson Air Force Base.

APPENDIX A

ORTHOGONAL INCOMPLETE FACTORIAL DESIGNS AND THEIR CONSTRUCTION USING LINEAR PROGRAMMING

SUMMARY

A necessary and sufficient condition is derived for an incomplete factorial design to be orthogonal under a very wide class of parameterizations. The condition is that for every subset of t factors, where t is twice the number of factors in the highest-order interaction to be estimated, each combination of levels occurs the same number of times. This condition is an extension of one due to Plackett (1946). From this condition can be derived a set of linear constraints on the number of occurrences of each treatment combination. The problem of finding the smallest orthogonal design to fit a given experimental situation is therefore an integer linear programming problem.

It is shown that designs can be derived using ordinary linear programming algorithms with a few special devices, rather than the more complicated integer linear programming techniques. As an example, a one-half replicate of the 2^4 is derived in detail.

It is desirable to have available from an experiment an estimate of experimental error obtained from duplicate runs. Such an estimate is unbiased by high-order interactions. It is shown how the linear programming technique can be modified to find the smallest orthogonal design containing at least one duplicated run.

1. INTRODUCTION

In this paper we consider the general class of orthogonal incomplete factorial designs. Any design which does not contain the full factorial will be considered as an incomplete factorial design, so that the class includes designs having more than one run at some of the treatment combinations. The class of orthogonal incomplete factorial designs contains as a subclass the familiar fractional factorials, for which the treatment combinations at the runs to be made form a subgroup or coset under a suitably defined group operation.

The definitions of such fundamental concepts as the parameterization underlying the analysis of factorial designs, aliasing in incomplete designs, and the resolution of a design are often defined in terms which apply only to fractional factorials. Since the discussion will not be limited to such designs, definitions which apply to any incomplete factorial design will be formulated.

Parameterization

The parameters for the model assumed in a factorial-design situation are often defined in terms of their estimators in the full factorial. For example, the main effect of a two-level factor may be defined as one-half the difference of the average response at the high

level and the average response at the low level of the factor (see, for example, Yates, 1935). For our purposes it is more convenient to define parameters directly in terms of the expected values of the responses. For notational simplicity we will define the parameters for a design containing only four factors. We will use s_r to designate the number of levels of the r-th factor, the integers $0, \cdots, s_r-1$ to designate the levels of each factor, and a set of integers in juxtaposition to designate a treatment combination. Let the expected value of the response for runs at treatment combination $ijk\ell$ be denoted by $\mu_{ijk\ell}$. A dot replacing a subscript will mean that the expected value has been computed over all values for that subscript. The grand mean is defined to be $\mu_{....}$. The main-effect parameters for the first factor are a set of s_1-1 linearly independent contrasts (linear combinations the sum of whose coefficients is zero) among the $\mu_{i...}$, and main effects for the other factors are defined analogously. The s_1-1 contrasts are usually taken to be orthogonal as well as linearly independent. The two-factor interaction parameters for the first two factors are a set of $(s_1-1)(s_2-1)$ contrasts among the quantities $\mu_{ij..}$, and these contrasts are usually taken to be orthogonal to each other as well as to the contrasts defining main effects. The definitions of the interaction parameters among other pairs of factors are analogous.

Similarly three- and four-factor interactions can be defined as contrasts among certain of the $\mu_{ijk\ell}$. If μ is the vector of $\mu_{ijk\ell}$ and if β^* is the vector of $\sum s_r$ parameters which have been defined, we may write $\beta^* = Q^{*-1}\mu$, or $\mu = Q^*\beta^*$. The matrix Q^* is in fact nonsingular from the given definition of β^* in terms of μ.

The effect of ignoring high-order interactions

Experience indicates that in most factorial situations the high-order interactions are so small that their estimates from an experiment are not likely to be statistically significantly different from zero. If the high-order interactions are ignored (treated as if they were zero) the model can be rewritten $\mu = Q\beta$, where β is a vector containing those components of β^* corresponding to non-ignored parameters, and where Q contains a corresponding selection of the columns of Q^*. If many components of β^* can be ignored, it may be advantageous to run an incomplete factorial design.

Let Y be a vector of observations from an incomplete factorial design, and suppose EY can be expressed as $EY = X\beta$ under the assumption that ignored parameters are zero, where X contains a selection of the rows of Q. Each column of X is associated with one component of β, and the column associated with a parameter will be called the coefficient vector for that parameter.

If the cross-product matrix $X'X$ is nonsingular, then $\hat{\beta} = (X'X)^{-1}X'Y$ is the least-squares estimate of β. If $X'X$ is nonsingular, we will say that β is measurable (Connor and Zelen, 1959, p. 2). Note that we have not actually assumed that β is estimable. When ignored parameters have the value zero, measurability implies estimability. The expected value of $\hat{\beta}$ is $(X'X)^{-1}X'X^*\beta^*$ (Box and Wilson, 1951, Appendix B), where X^* consists of rows of Q^*, so that the expected value of a component $\hat{\beta}_0$ of $\hat{\beta}$ is β_0 plus a linear combination of ignored parameters. This effect is called <u>aliasing</u> (not confounding) and β_0 is said to be aliased with the linear combination of ignored parameters. Confounding is a special case of aliasing in which block contrasts are aliased with ignored interaction parameters.

The concept of the <u>resolution</u> of a design, introduced by Box and Hunter (1961), will prove useful, and a slightly modified definition is as follows. If a design is such that all parameters involving r or fewer factors are measurable ignoring all interactions of $r+1$ or more factors, the design is said to be of resolution $2r+1$; if all parameters involving $r-1$ or fewer factors are measurable ignoring all interactions of $r+1$ or more factors, the design is said to be of resolution $2r$.

Conditions for optimality of orthogonality

It has often been stated that orthogonal designs yield the maximum possible information per run for each parameter estimate. The actual theorem was first proved by Plackett and Burman (1946). If a class of designs is considered for which the lengths of the column vectors of the matrix X are fixed, and if there is an orthogonal design in this class, then the variances of all estimates are individually minimized over the class using the orthogonal design.

The condition that the lengths of the columns be fixed does more than fix the scale of measurement, however. If the parameterization is fixed it often happens that, for some of the parameters, non-orthogonal designs have coefficient vectors of greater length than those for orthogonal designs with the same number of runs. It can also happen that non-orthogonal designs have lower variances for some of the estimates than orthogonal designs with the same number of runs as in the following example.

Consider an experimental situation with a single factor at three levels. Let μ_0, μ_1, and μ_2 be the expected value of the response at the three levels, and let the parameters to be estimated be the grand mean, $(\mu_0 + \mu_1 + \mu_2)/3$, the "linear effect", $(\mu_2 - \mu_0)/2$, and the "quadratic effect", $(\mu_0 - 2\mu_1 + \mu_2)/6$. The design consisting of

seven runs at each level is orthogonal, and the variances of the estimates are σ^2 times 1/21, 1/14, and 7/294, respectively. The design consisting of eight runs at levels 0 and 2 and five runs at level 1 is non-orthogonal and has variances σ^2 times 1/20, 1/16, and 7/240. The fact that non-orthogonal designs may be superior in this sense to orthogonal ones has apparently never been utilized in the construction of incomplete factorial designs.

Although orthogonal designs are not in general optimum in the sense of minimizing all the variances, experience indicates that they are rarely far from optimum. In the important case of factors at two levels, orthogonal designs are optimum because all coefficient vectors consist of plus and minus ones, so that the lengths of the columns of X are automatically fixed. It has previously been shown (Webb, 1964) that for most parameterizations, orthogonal designs satisfy other criteria for optimality (see Section 2 below). In addition, they afford great computational advantages, and therefore there is still reason to be very much concerned with methods for constructing orthogonal designs.

In a subsequent paper it will be shown how linear constraints can be used to characterize classes of non-orthogonal designs.

2. THE CHARACTERIZATION OF ORTHOGONALITY

Tensor product parameterizations

A design is orthogonal if $X'X$ is diagonal, so that orthogonality of a design depends on the parameterization used to describe the response. A characterization of orthogonality will be derived first for a special class of parameterizations which will be referred to as tensor product parameterizations and then extended to a larger class of parameterizations.

Let U and V be m and n dimensional vector spaces with bases $u_1, \cdots, u_m$ and $v_1, \cdots, v_n$, respectively. Let w_{ij} be an mn dimensional vector whose components are the mn possible products taken in a given order of a component of u_i with a component of v_j. As i goes from 1 to m and j goes from 1 to n, the mn vectors which are produced in this manner are linearly independent. Let W be the vector space spanned by the w_{ij}. Then W is called the tensor product of the vector spaces U and V, and the basis $\{w_{ij}\}$ is called the product basis of the component bases $\{u_i\}$ and $\{v_j\}$. If the bases $\{u_i\}$ and $\{v_j\}$ are orthogonal, then the basis $\{w_{ij}\}$ is an orthogonal basis for the tensor product.

Suppose that in an experiment involving n factors the main effects of the i-th factor are defined as orthogonal contrasts among the expected values averaged over the levels of the other factors. The coefficient vectors of these $s_i - 1$ contrasts, together with the s_i-component vector all of whose components are the constant $1/s_i$, form an orthogonal basis for a vector space V_i. Let the two-factor interaction parameters between the i-th and j-th factors be defined by all contrasts (except those which define main effects and the grand mean) whose coefficient vectors are in the product basis of the tensor product of V_i and V_j. The contrasts are between the expected responses averaged over the levels of the remaining $n-2$ factors. Similarly, three-factor through n-factor interactions are defined in terms of the tensor product of sets of the V_i.

Since the basis for each V_i is orthogonal, the product bases for tensor products of sets of the V_i are also orthogonal. Therefore, the rows of the matrix Q^{*-1} appearing in the relationship $\beta^* = Q^{*-1}\mu$ are mutually orthogonal. A column of Q^* is therefore a multiple of the corresponding row of Q^{*-1}. Since the components of the row of Q^{*-1} defining the grand mean are all equal to the constant $1/\Pi_{i=1}^{n} s_i$, and since this matrix is of order $\Pi_{i=1}^{n} s_i$, the coefficient vector for the grand mean in any (complete or incomplete) factorial design has all its components equal to unity. Due to

the tensor product parameterization, a component of the coefficient vector for an interaction parameter is the product of the corresponding components of the coefficient vectors for main effects of factors making up the interaction.

It should be noted that such parameterizations are not new. In particular, the equivalent of tensor product parameterizations was used by Plackett (1946, p. 330) in deriving conditions for orthogonality, although he did not use our present terminology.

The characterization

Plackett (1946) was concerned with the situation in which the main effects of a number of factors are to be estimated together with the interaction between any two of them. He found that in order that the interaction parameters between A and B be estimated orthogonally, all combinations of levels of A and B must appear equally often with the levels of every other factor. In this case the interactions between A and every other factor, and between B and any other factor, may also be estimated orthogonally. In the present paper we are concerned with the situation in which all interactions of a given order are to be estimated. It is seen, therefore, that Plackett's result does not apply. Our general characterization is given in Theorem 4, the first three theorems giving preliminary results.

THEOREM 1. If in a given design the i-th factor appears at s_i levels, the coefficient vectors for all its main-effect parameters are orthogonal to the coefficient vector for the grand mean if and only if every level of the factor appears the same number of times.

PROOF. Let the components of the s_i-component vector z be the numbers of occurrences of each level of the factor in the design. The inner product of the coefficient vector for a main-effect parameter and the coefficient vector for the grand mean is equal to a non-zero scalar times the inner product of z and the vector in V_i which defines the main effect. The vector z is orthogonal to all the vectors in V_i defining main effects if and only if z is a multiple of the remaining basis vector, namely the vector of constants. Therefore, all components of z are equal. This concludes the proof.

We now ask when all the coefficient vectors for main effects of one factor are orthogonal to all the coefficient vectors for main effects of another factor. Because the tensor product parameterization is being used, the inner product of two coefficient vectors, one for a main effect of one factor and the other for a main effect of the other, is numerically the same as the inner product of the coefficient vectors for one of the two-factor interaction parameters and the grand mean. We will now prove that this observation is true more generally.

THEOREM 2. All coefficient vectors associated with the m-factor interactions among factors $A_1, \cdots, A_m$ are orthogonal to those for the n-factor interactions among factors $B_1, \cdots, B_n$ if and only if all coefficient vectors for the (m+n)-factor interactions are orthogonal to the coefficient vector for the grand mean.

PROOF. Because of the way interactions were defined, there is an (m+n)-factor interaction parameter whose coefficient vector has as components the product of coefficients for any given "A-interaction" and any given "B-interaction". Therefore, the inner product of coefficient vectors for an A-interaction and a B-interaction is numerically equal to the inner product of coefficient vectors for an "AB-interaction" and the grand mean.

Theorem 1 can now be generalized to include orthogonal designs for estimating all the effects and interactions among several factors.

THEOREM 3. Given ν factors $A_1, \cdots, A_\nu$, the coefficient vectors for the grand mean, main effects, and interactions of all orders are mutually orthogonal if and only if all possible combinations of levels appear the same number of times.

PROOF. If every combination appears the same number of times then the design is a replicated full factorial in the ν factors

$A_1, \cdots, A_\nu$ so that in particular the design is orthogonal. Now suppose all the coefficient vectors are orthogonal. In particular, all the coefficient vectors for main effects and interactions are orthogonal to that for the mean. In the space which is the tensor product of V_1 through V_ν, let z be the vector whose components specify how many times each treatment combination appears in the design. Any given effect or interaction parameter is defined in terms of some vector, call it b, which is also in the tensor product. Then the inner product of the coefficient vector for the given parameter with the coefficient vector for the grand mean is a scalar times $z'b$. The vector z is required to be orthogonal to all such b; therefore z must again have all its components equal, which concludes the proof.

In view of Theorems 2 and 3, we have the following immediate corollary, which serves to characterize orthogonal resolution $t+1$ designs.

THEOREM 4. A design of resolution $t+1$ is orthogonal under any parameterization derived from a complete tensor product parameterization by ignoring interactions involving $\frac{1}{2}(t+1)$ or more factors if and only if for every subset of t factors all possible combinations of levels appear the same number of times.

PROOF. If t is even, all coefficient vectors for parameters involving $\frac{1}{2}t$ factors are required to be orthogonal to coefficient vectors for other parameters involving $\frac{1}{2}t$ or fewer factors. By Theorem 2, this happens if and only if all coefficient vectors for parameters involving t or fewer factors are orthogonal to the coefficient vector for the grand mean. By Theorem 3, this is equivalent to the requirement that for every subset of t factors all possible combinations of levels appear the same number of times. If t is odd, all coefficient vectors for parameters involving $\frac{1}{2}(t-1)$ factors are required to be orthogonal to coefficient vectors for other parameters involving $\frac{1}{2}(t+1)$ or fewer factors, which again is the case if and only if all coefficient vectors for parameters involving t or fewer factors are orthogonal to the coefficient vector for the grand mean. The remainder of the proof proceeds as when t is even.

Alternative parameterizations

Theorem 4 remains true under a wide variety of parameterizations. In fact, we can state the following:

THEOREM 5. If β is a vector of parameters which are defined by ignoring interactions of $r+1$ or more factors from a complete tensor product parameterization β^*, if α is a linear transform of β, say $\alpha = C^{-1}\beta$, and if the full factorial design is orthogonal under

the parameterization α, then Theorem 4 is valid for the parameterization α.

PROOF. Let Q and X be the coefficient matrices under the parameterization β for the full factorial design and for any other orthogonal design with say N runs, respectively. Then the conditions of the theorem imply that $Q'Q$, $X'X$, and $C'Q'QC$ are diagonal matrices. But since X and Q represent designs orthogonal under the same tensor product parameterization, $X'X$ must be equal to the constant $N/\Pi_i s_i$ times $Q'Q$. Therefore, $C'X'XC$ is also diagonal, as was to be proved.

Recently, a theorem has been published by Addelman (1962) which states that a design of resolution 3 is orthogonal if and only if for every pair of factors the number of occurrences of the combination of levels ij is given by $n_{ij} = n_{i.}n_{.j}/N$. Here $n_{i.}$ and $n_{.j}$ are the numbers of occurrences of level i of the first factor and level j of the second factor, respectively. Such designs are called proportional-frequency designs. The apparent contradiction between Addelman's theorem and Theorem 4 above stems from a different definition of grand mean. His grand mean is the expected value of the average of the observations made, while the one used above is the average of the expected values at all possible treatment combinations.

Although the first published proof of this theorem is apparently that of Addelman (Addelman and Kempthorne, 1961), its validity was apparently recognized by Stevens (1948). Using the present definition of the grand mean, but defining main effects by linear combinations of the $\mu_{i...}$ other than orthogonal contrasts, Plackett (1946) showed that the condition of proportional frequencies is necessary, but not sufficient, for a design to be orthogonal.

The question naturally arises as to which definition of grand mean is preferable. If it can be supposed that the mean is only a nuisance parameter, as for example in a screening experiment, Addelman's definition would likely be the more useful. If, however, the purpose of the experiment is to describe the response over the points of the full factorial, then the present definition seems more appropriate as can be seen by the following argument.

Since we are supposing that all points are of interest, the criterion for judging designs should depend on the variances of predicted values at all treatment combinations of the full factorial. Since these variances approach zero as the number of runs increases, the criterion should be adjusted in some way for the number of runs, N. A convenient and realistic criterion is therefore the average variance of predicted values multiplied by N. Under this criterion the best

main-effect designs are those which are orthogonal using the present definition of the grand mean, as the following theorem demonstrates. The theorem was given previously (Webb, 1964), and is reproduced here for convenience.

THEOREM 6. In a design for estimating parameters under any main-effect parameterization, the average variance of a predicted value times the number of runs is minimized if and only if the design is an equal-frequency design. In this case the variances of all predicted values are equal.

PROOF. Let X be the coefficient matrix of a design under some main-effect parameterization. Let Z be the coefficient matrix of the full factorial. Let X contain N rows, Z contain M rows, and let there be p parameters. The value of the criterion being considered is $(\sigma^2 N/M)\text{trace } Z(X'X)^{-1}Z'$, where σ^2 is the error variance of an observation.

If $EY = X\beta$ under the given parameterization and A is any nonsingular matrix, then under the parameterization $\alpha = A^{-1}\beta$ the covariance matrix of predicted values is $\sigma^2 ZA(A'X'XA)^{-1}A'Z'$, which equals $\sigma^2 Z(X'X)^{-1}Z'$. Therefore the choice of parameterization is arbitrary. Without loss of generality we will choose a tensor product parameterization in which the main effects are scaled so that $Z'Z$ is equal to the constant M times the identity matrix.

Now choose any factor at say s levels and consider columns of Z corresponding to the $s-1$ main effects of the given factor. By rearranging the rows of Z, these columns can be expressed as an s by $s-1$ matrix S repeated vertically M/s times. Letting i be a vector containing s ones, we observe now that the matrix $1/\sqrt{s} \cdot [i,S]$ is orthonormal. It follows that the sum of squares of any row (or column) is equal to unity. Therefore the sum of squares of any row of S is equal to $s-1$. This argument may be repeated for each factor, so that the sum of squares of any row of Z is equal to p, the number of parameters. Since the matrix X contains a selection of the rows of Z, the sum of squares of any row of X is also p.

Returning to the average variance of a predicted value, note that $\sigma^2/M \text{ tr } Z(X'X)^{-1}Z' = \sigma^2/M \text{ tr } (X'X)^{-1}Z'Z = \sigma^2 \text{ tr } (X'X)^{-1}$. But the last expression is equal to $\sigma^2 \cdot \Sigma_{i=1}^{p} 1/\lambda_i$, where the λ_i are the eigenvalues of $X'X$. However, the sum $\Sigma \lambda_i$ is fixed since $\Sigma \lambda_i = \text{tr } X'X = \text{tr } XX' = N \cdot p$ (the sum of squares of any row of X is equal to p). If the sum of p positive numbers if fixed, the sum of the reciprocals is minimized if the numbers are all equal. The eigenvalues of a matrix are equal if and only if the matrix is a constant times the identity. Therefore the average variance of a predicted value is minimized if and only if $X'X$ is equal to N times

the identity and then the average variance is $\sigma^2 p/N$. Since a tensor product parameterization was used, Theorem 4 applies and the criterion attains its minimum value $\sigma^2 p$ whenever for each subset of two factors each combination of levels appears the same number of times. Furthermore this result is independent of the parameterization used and of the number of runs. If $X'X$ is equal to a constant times the identity matrix it is easy to see that the diagonal elements of $Z(X'X)^{-1}Z'$ are all equal. This completes the proof of Theorem 6.

3. CONSTRUCTING DESIGNS BY LINEAR PROGRAMMING

Conversion to a linear programming problem

In this section a detailed method for recharacterizing an orthogonal design as an integer solution to a linear programming problem is presented. First, consider the special case of symmetric designs, in which all n factors are at s levels. Theorem 4 implies that, for any t factors, each combination of levels occurs the same number of times, λ. This number is called the index of the design (Bose and Bush, 1952), and the number of runs, N, in the design is given by $N = \lambda s^t$. Let J_0 be a collection of all treatment combinations which are identical in t components, say the first t. There will be s^{n-t} vectors in J_0. Let w_j be the number of occurrences of

treatment combination j in an orthogonal resolution $t+1$ design. Then Theorem 4 requires that $\Sigma_{j\epsilon J_0} w_j$ be equal to λ. But the first t components may be fixed at any of s^t sets of values, each of which yields a constraint $\Sigma_{j\epsilon J} w_j = \lambda$. This argument may be repeated for all $\binom{n}{t}$ subsets of t factors, so that in all $s^t \times \binom{n}{t}$ constraints may be written down.

The analogous results are more complicated for asymmetric designs, in which the n factors appear at $s_1, s_2, \cdots, s_n$ levels, respectively. For convenience, we first extend the definition of index to apply to asymmetric designs as well.

Let I_γ be a subset of t of the integers 1 through n, and let the collection $\{I_\gamma | \gamma \epsilon \Gamma\}$ contain every such subset. The index class Γ will be the integers $1, 2, \cdots, \binom{n}{t}$. Let K be the least common multiple of the numbers $\Pi_{i\epsilon I_\gamma} s_i$, as γ ranges over Γ. The number of runs in an orthogonal resolution $t+1$ design must be a multiple of K, say $N = \lambda K$. This number λ may be defined as the index of the design.

Suppose J_γ is one of the $\Pi_{i\epsilon I_\gamma} s_i$ collections of all those treatment combinations j which have fixed values for the set of components whose ordinal numbers are in the set I_γ. Then Theorem 4 implies $\Sigma_{j\epsilon J_\gamma} w_j = K\lambda / \Pi_{i\epsilon I_\gamma} s_i$. As γ ranges through Γ, a set of

constraints is generated which must be satisfied by any orthogonal design. This set is linearly dependent. Two methods for reducing the complete set to an equivalent set of linearly independent constraints are described in the Appendix. The constraint matrices given in the examples below were constructed by using one of these methods (the "base-set method").

Eliminating unacceptable solutions

The problem of finding the orthogonal design with the smallest value of N for a given set of factors and given resolution is now equivalent to finding a set of non-negative integer values for the variables w_j and λ such that the set of constraints summarized by $\Sigma_{j\in J_\gamma} w_j = K\lambda/\Pi_{i\in I_\gamma} s_i$ is satisfied and such that λ is minimized. If the problem is stated in this form, then $w_j = 0$, $\lambda = 0$ provides a trivial solution. To rule out this solution, the additional requirement that λ be positive can be added. A convenient way to include this requirement is to add a new non-negative variable λ^* and the constraint $\lambda - \lambda^* = 1$.

The full factorial is orthogonal and hence satisfies the constraints with each variable w_j equal to unity and the number of runs equal to $\Pi_{i=1}^{n} s_i$. By multiplying each variable by the fraction $K/\Pi s_i$ another unacceptable solution is obtained with λ equal to one. It

follows from Theorem 4 that an orthogonal design remains orthogonal if the designations of the levels of any factor are rearranged. Since a design must have at least one run at some treatment combination, we may require that an arbitrary one of the w_j, say $w_{0\ldots0}$, be greater than or equal to one. The addition to the problem of a non-negative variable $w^*_{0\ldots0}$ and the constraint $w_{0\ldots0} - w^*_{0\ldots0} = 1$ will rule out the unacceptable solution corresponding to $\lambda = 1$.

Non-integer solutions may still exist, but in the examples worked to date, the use of the simplex method for linear programming together with a simple algorithmic device has led to the smallest orthogonal designs. The device is as follows: if the minimum solution to the problem obtained by the simplex method is such that λ is not an integer, change the problem by requiring that λ be greater than or equal to the next larger integer. This can be done by revising the constraint $\lambda-\lambda^* = 1$, rather than by adding a new constraint, so that some of the previous feasible solutions may still be feasible solutions to the revised problem. Thus, it is not necessary to start the linear programming procedure over from the beginning.

Numerical example

The following example is intended only to illustrate the steps involved in constructing designs by linear programming, the resulting

design being a familiar fractional factorial. In Section 4 some designs will be derived which are not fractional factorials, but their construction does not illustrate the general procedure. It will be assumed that the reader is familiar with the basic simplex method as described in any linear programming text, for example, Hadley (1962).

We will derive the smallest design of resolution 3 for a 2^4 experimental situation. For convenience we abandon the vector subscripts on the w_j and substitute numerical subscripts. Let w_j be the number of occurrences of that treatment combination which is the binary expansion of the decimal number j. Thus w_0 is the number of occurrences of 0000, w_1 is the number of occurrences of 0001, and so forth. The independent linear constraints (obtained by using the base-set method described in the Appendix) are given by $Aw = b$, where $w = (\lambda, \lambda^*, w_0^*, w_0, w_1, \cdots, w_{15})'$, $b = (1, 1, 0, \cdots, 0)'$, and A is the following matrix (a dot is used for zero).

$$
\begin{bmatrix}
1 & -1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & . & -1 & 1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & -1 & -1 & -1 & -1 & . & . & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & -1 & -1 & . & . & -1 & -1 & . & . & . & . & . & . & . & . & . & . \\
1 & . & . & -1 & . & -1 & . & -1 & . & -1 & . & . & . & . & . & . & . & . & . \\
1 & . & . & . & . & . & . & -1 & -1 & -1 & -1 & . & . & . & . & . & . & . & . \\
1 & . & . & -1 & -1 & . & . & . & . & . & . & -1 & -1 & . & . & . & . & . & . \\
1 & . & . & -1 & . & -1 & . & . & . & . & . & -1 & . & -1 & . & . & . & . & . \\
1 & . & . & . & . & . & . & . & . & . & . & -1 & -1 & -1 & -1 & . & . & . & . \\
1 & . & . & -1 & . & . & . & -1 & . & . & . & -1 & . & . & . & -1 & . & . & . \\
1 & . & . & . & . & . & . & . & . & . & . & -1 & -1 & . & . & -1 & -1 & . & . \\
1 & . & . & . & . & . & . & . & . & . & . & -1 & . & -1 & . & -1 & . & -1 & . \\
1 & . & . & . & . & . & . & . & . & . & . & . & . & . & . & -1 & -1 & -1 & -1 \\
\lambda & \lambda^* & A_0^* & A_0 & A_1 & A_2 & A_3 & A_4 & A_5 & A_6 & A_7 & A_8 & A_9 & A_{10} & A_{11} & A_{12} & A_{13} & A_{14} & A_{15}
\end{bmatrix}
$$

In order to start the simplex method, an initial basic feasible solution is required. The solution corresponding to the full factorial design is $\lambda = 4$, $\lambda^* = 3$, $w_0^* = 0$, and $w_j = 1$ for $j = 0, 1, \cdots, 15$. This solution is feasible but not basic.

Denoting the columns of the matrix A by Λ, Λ^*, A_0^*, A_0, A_1, $\cdots$, A_{15}, and substituting the full-factorial solution into the equation $Aw = b$, we obtain the equation $4\Lambda + 3\Lambda^* + \Sigma_{j=0}^{15} A_j = b$. An obvious linear relationship among the columns of A is $\Lambda + \Lambda^* + A_3 + A_5 + A_6 - A_7 + A_8 + A_{15} = 0$. Subtracting this equation from the previous one yields the equation $3\Lambda+2\Lambda^*+A_0+A_1+A_2+A_4+2A_7+A_9+A_{10}+A_{11}+A_{12}+A_{13}+A_{14} = b$, to which corresponds the solution $\lambda = 3$, $\lambda^* = 2$, $w_7 = 2$, $w_j = 1$ for $j = 0, 1, 2, 4, 9, 10, 11, 12, 13$, and 14, and the remaining variables zero. It can be verified that the columns with non-zero coefficients in the last equation are linearly independent. Moreover, there are exactly thirteen such columns, the same as the number of rows, so that this solution is basic and feasible. For other designs a set of linearly independent columns of the constraint matrix can be obtained by using the solution corresponding to a known fractional factorial. It may then be necessary to adjoin additional columns in order to have a full basis.

Returning to the present example, let B be the basis matrix consisting of the chosen set of linearly independent columns, that is,

$B = [\Lambda, \Lambda^*, A_0, A_1, A_2, A_4, A_7, A_9, A_{10}, A_{11}, A_{12}, A_{13}, A_{14}]$. In applying the simplex method we first compute $G = B^{-1}A$ and $w_B = B^{-1}b$. Those columns of G which are not columns of the identity matrix and the vector w_B are as follows:

$$[G_0^*, G_3, G_5, G_6, G_8, G_{15}] = \begin{bmatrix} -3 & 1 & 1 & 1 & -2 & -2 \\ -3 & 1 & 1 & 1 & -2 & -2 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & -1 & -1 \\ -1 & 1 & 1 & 0 & -1 & -1 \\ -1 & 0 & 1 & 1 & -1 & -1 \\ -2 & 1 & 1 & 1 & -1 & -1 \\ -1 & 0 & 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & -1 \\ -1 & 1 & 0 & 0 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 & -1 \\ -1 & 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}, \quad w_B = \begin{bmatrix} 3 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Since the linear form to be minimized is just the first variable λ, a column of A should be introduced into the basis only if the first component of the corresponding column of G is positive, for only then could λ decrease. The column A_3 will be chosen to enter. In order to maintain feasibility, the column chosen to leave the basis must be such that for $g_{i3} > 0$ (g_{i3} is the i-th component of column G_3) the ratio w_{Bi}/g_{i3} is minimized. The smallest value is 1 which is attained for $i = 4$, 5, 10, and 11. If we allow A_1 to leave the basis (for which $i = 4$) and recompute G and w_B using the revised basis matrix, we obtain:

$$[G_0^*, G_1, G_5, G_6, G_8, G_{15}] = \begin{bmatrix} -2 & -1 & 0 & 1 & -1 & -1 \\ -2 & -1 & 0 & 1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 & -1 & -1 \\ 0 & -1 & -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 & 0 & -1 \\ 0 & -1 & -1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & -1 & 0 \end{bmatrix}, \quad w_B = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

The solution now corresponds to one of the half-replicates of a 2^4 design. Since G_6 is the only one of these vectors whose first component is positive, A_6 is the only vector whose admission into the basis could make λ smaller. For the fifth row the ratio w_{Bi}/g_{i6} has the value zero, so that the corresponding column of A, namely A_2, is the only column which can leave. Since the minimum value is zero, the solution does not change. The basis matrix is now $B = [\Lambda, \Lambda^*, A_0, A_3, A_6, A_4, A_7, A_9, A_{10}, A_{11}, A_{12}, A_{13}, A_{14}]$.

The columns of G of interest are now the following:

$$[G_0^*, G_1, G_2, G_5, G_8, G_{15}] = \begin{bmatrix} -2 & 0 & -1 & 1 & -1 & -1 \\ -1 & 0 & -1 & 1 & -1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & -1 & -1 \\ 0 & -1 & 1 & -1 & 0 & 0 \\ -1 & 1 & -1 & 2 & -1 & -1 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 1 & -1 & 1 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 & -1 & 0 \\ -1 & 1 & -1 & 1 & -1 & 0 \end{bmatrix}, \quad w_B = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

The vector A_5 can be introduced into the basis, and the minimum value of the ratio w_{Bi}/g_{15} is $\frac{1}{2}$, so that A_4 must leave the basis. The new solution is $w = (1\frac{1}{2}, \frac{1}{2}, 1, 0, 0, 0, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0)'$. By recomputation of G it may be verified that no new column can come into the basis, so that the present solution is a minimum basic feasible solution. It does not consist of integers, however, and therefore does not correspond to an actual design. Nevertheless, we now know that there can be no design with $\lambda = 1$, so the smallest design must have $\lambda \geq 2$. We have incidently obtained a solution with $\lambda = 2$, however, so that the half-replicate corresponding to that solution is one of the class of smallest orthogonal designs, as are the other known half-replicates of a 2^4 design.

4. ALTERNATIVE ORTHOGONAL DESIGNS AND PARTIAL DUPLICATION

The technique of viewing an experimental design as a solution to a set of linear constraints lends itself naturally to the investigation of possible alternative designs. Consider a simple experiment involving two factors at two levels and one at three levels, and suppose that only the grand mean and main effects are to be estimated. For this experimental situation the constant K is equal to twelve, and therefore an orthogonal design must contain 12λ runs, where λ is the index. The full factorial is orthogonal and involves exactly twelve runs, and therefore is a smallest orthogonal resolution 3 design for this situation. One may ask whether there are any other orthogonal designs with just twelve runs.

Let the third component of the treatment combination vector represent the level of the three-level factor. The linear constraints may be written in a form expressing each w_j as a linear combination of λ, w_0, and w_1, where w_0 and w_1 are the numbers of occurrences of treatment combinations 000 and 001. These linear combinations appear in the second column of the following table.

Treatment Combination	Variable	Solutions		
000	w_0	1	1	2
001	w_1	1	0	1
002	$w_2 = 3\lambda - w_0 - w_1$	1	2	0
010	$w_3 = 2\lambda - w_0$	1	1	0
011	$w_4 = 2\lambda - w_1$	1	2	1
012	$w_5 = w_0 + w_1 - \lambda$	1	0	2
100	$w_6 = 2\lambda - w_0$	1	1	0
101	$w_7 = 2\lambda - w_1$	1	2	1
102	$w_8 = w_0 + w_1 - \lambda$	1	0	2
110	$w_9 = w_0$	1	1	2
111	$w_{10} = w_1$	1	0	1
112	$w_{11} = 3\lambda - w_0 - w_1$	1	2	0
	λ	1	1	1

The full factorial corresponds to taking $w_0 = 1$, $w_1 = 1$, and $\lambda = 1$, and this is the first of the three solutions given in the table.

If w_0 and w_1 are assigned values it is possible to construct a design with these values if λ is chosen properly. It seems plausible that if the value of w_1 is decreased to zero, then the value of λ would not have to increase. This is indeed the case, and the corresponding solution is the second one given in the above table. The structure of this design is interesting. If the 0 level of the third factor is ignored, the design is a half-replicate of a 2^3 run

twice. Also given in the above table is a third design, which is a repeated half-replicate of 2^3 if level 1 of the third variable is ignored.

There are several alternative designs available in many experimental situations. In the case of a $2 \cdot 3^2$ situation there are, in addition to the full factorial, orthogonal resolution 3 designs with the same number of runs in which four treatment combinations are duplicated and others in which six treatment combinations are duplicated.

There are two basic reasons for the selection of alternative designs. The first is that sometimes certain sets of experimental conditions can not be attained, and there may be an alternative design which does not include the taboo treatment combinations. The second reason is that alternative designs often provide partial duplication.

The argument for partial duplication has been given by Daniel (1957) and by Dykstra (1959). Briefly, it is this: if partial duplication is present, an error estimate is available which is unaffected by the presence of high-order interactions. Dykstra gave a catalogue of partially duplicated designs, some of which are due to Daniel. The designs catalogued are constructed by combining pairs of fractional factorial designs, and are all non-orthogonal.

The linear programming procedure for orthogonal designs can be modified to obtain a procedure for finding the smallest orthogonal design involving partial duplication. Rather than requiring that w_0 be greater than or equal to one, we require that w_0 be greater than or equal to two. Consider again a 2^4 experimental situation as in the numerical example of Section 3. The constraint matrix A remains unchanged. Since the smallest orthogonal design has λ equal to 2, we may start with the constraint $\lambda-\lambda^* = 2$. The vector b has its first two components equal to 2 and the remaining components equal to zero. The minimum basic feasible solution to this revised problem is twice that which was obtained in the example; that is, $w = (3, 1, 0, 2, 0, 0, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0)'$. Among incomplete 2^4 designs of resolution 3, the corresponding design is therefore the smallest which is orthogonal and involves partial duplication. A single duplicated point would ordinarily not be enough, and since the given design is a unique minimum, a larger design is required. In sixteen runs the half-replicate can be repeated, so that in practical situations one would either use this design or a non-orthogonal design such as one of those given by Dykstra.

APPENDIX

In Section 3 it was noted that the set of all possible constraints implied by Theorem 4 is linearly dependent. In this Appendix are presented two methods for arriving at equivalent sets of linearly independent constraints for orthogonal factorial designs. These will be called the base-set method and the crossing-out method. The base-set method consists of finding a base set of the w_j and λ, which is any largest linearly independent subset of the collection of w_j and λ. Once a base set is found, all the constraints can be written down as expressions for w_j not in the base set. The crossing-out method involves writing down all possible constraints and then systematically crossing out those which are linear combinations of previous ones.

This Appendix contains details of the two techniques, examples of their use, and a proof of their validity. The proof will proceed by showing that the constraints arrived at by using the base-set method are linearly independent, that the constraints arrived at by using the crossing-out method are equivalent to the set of all constraints, and finally that the same number of constraints is obtained by using either method.

The base-set method

Once again let w_j represent the number of occurrences of that treatment combination whose vector representation is the vector j. That level of a factor which is designated by the largest number will be called the "highest level" for that factor. We will see that the set of w_j for which j has n-t-1 or fewer components at their highest levels is a base set for an orthogonal resolution t+1 design for $n \geq t$ factors. Obviously this is but one of a large number of base sets. If n equals t, then every w_j is equal to λ, so that the base set consists of just this element.

For convenience, the alleged base set will be referred to as the base set, even though this will not be established until later. If it can be shown that each w_j not in the base set can be expressed as a linear combination of variables in the base set, then it follows that the set of constraints generated in this way is linearly independent, since each involves a unique variable not appearing in any other constraint. It is enough to show that an expression for each w_j not in the base set can be derived in terms of λ and w_j with fewer components of j at their highest levels.

Let w_J be any w_j such that J has exactly n-t+h components at their highest levels, where h is an integer between 0 and t. Then w_J is not in the base set. For convenience, let these components

be the last $n-t+h$. By Theorem 4, the number of times the first t factors occur together at the levels appearing as the first t components of J is a known multiple M of λ. (The constant M depends on the numbers of levels of the n factors and on which are at their highest levels in J.) Therefore, $M\lambda$ equals the sum of all those w_j for which j is identical with J in the first t components. This constraint can be rewritten as w_J equals $M\lambda$ minus a sum of w_j for which j has $n-t+h-1$ or fewer components at their highest levels. Therefcre, as asserted, any w_j not in the base set can be expressed as a linear combination of λ and w_j with fewer components at their highest levels.

To illustrate we will construct a set of linearly independent constraints for the resolution-3, 2^4 design, as in the example of Section 3. A base set is $\{w_{0000}, w_{0001}, w_{0010}, w_{0100}, w_{1000}, \lambda\}$. By fixing the first two components of j at 0 and 0, we see that $w_{0000} + w_{0001} + w_{0010} + w_{0011} = \lambda$, which yields an expression for w_{0011} in terms of elements of the base set. The following table lists in schematic form expressions for the w_j not in the base set in terms of variables whose subscripts have fewer components at their high levels. The subscripts j are used for the variables w_j. The third column lists the ordinal numbers of the components which are held fixed in deriving the given expressions.

Variable	Expression	Fixed Components
0011	λ - 0000 - 0001 - 0010	1 and 2
0101	λ - 0000 - 0001 - 0100	1 and 3
0110	λ - 0000 - 0010 - 0100	1 and 4
0111	λ - 0100 - 0101 - 0110	1 and 2
1001	λ - 0000 - 0001 - 1000	2 and 3
1010	λ - 0000 - 0010 - 1000	2 and 4
1011	λ - 1000 - 1001 - 1010	1 and 2
1100	λ - 0000 - 0100 - 1000	3 and 4
1101	λ - 1000 - 1001 - 1100	1 and 3
1110	λ - 1000 - 1010 - 1100	1 and 4
1111	λ - 1100 - 1101 - 1110	1 and 2

This same set of constraints, written in matrix form, is used in the example.

The number of constraints using the base-set method is the same as the number of w_j not in the base set, and this in turn is the number of treatment combinations with $n-t$ or more components at their highest levels. Let t equal 2. The number of treatment combinations with n components at their highest levels is 1, the number with $n-1$ components at their highest levels is $\Sigma_{i_1=1}^{n} (s_{i_1}-1)$, and the number with $n-2$ components at their highest levels is $\Sigma_{i_1=1}^{n} \Sigma_{i_2=i_1+1}^{n} (s_{i_1}-1)(s_{i_2}-1)$. The total is therefore $1 + \Sigma(s_{i_1}-1) + \Sigma(s_{i_1}-1)(s_{i_2}-1)$. If t equals 3, the total number of constraints is $1+\Sigma(s_{i_1}-1)+\Sigma(s_{i_1}-1)(s_{i_2}-1)+\Sigma(s_{i_1}-1)(s_{i_2}-1)(s_{i_3}-1)$, where the range of summation in the last term is $1 \leq i_1 < i_2 < i_3 \leq n$.

The general expression can be written down in the form $1 + \Sigma_{k=1}^{t} \Sigma_{\theta_k} \Pi_{j=1}^{k} (s_{i_j} - 1)$, where the set θ_k is the set of i_j such that $1 \leq i_1 < i_2 < \cdots < i_k \leq n$.

The crossing-out method

The crossing-out method is a technique for reducing the set of all constraints by eliminating those which are obvious linear combinations of others. The set of all constraints contains statements about the number of occurrences of combinations of levels of subsets of t factors. By adding together sets of constraints, statements can be derived about the number of occurrences of combinations of levels of $t-1$ factors. Because a given subset of $t-1$ factors is contained in several subsets of t factors, such statements are not unique. Indeed, it is because of this that the set of all constraints is linearly dependent.

The crossing-out technique will first be described for $t = 2$. Consider the first two factors, which are at s_1 and s_2 levels, respectively. A set of $s_1 s_2$ constraints can be written down from the requirement that the number of occurrences of each of the $s_1 s_2$ possible combinations of levels of the first two factors must be equal to each other and to a known multiple M_1 of the index λ. Let the first two components of j be represented by u and v, and suppose

that the constraints are written down in increasing numerical order of uv. That is, the constraint that $M_1\lambda$ equals the sum of all w_j for which the first two components of j are u and v has ordinal number s_2u+v+1 in the constraint array. If the first s_2 relationships are added together, the result is that the number of occurrences of level 0 of the first factor is equal to $s_2M_1\lambda$. By adding any of s_1 sets of s_2 consecutive constraints, the number of occurrences of any level of the first factor is $s_2M_1\lambda$ also. Similarly, the number of occurrences of any level of the second factor equals $s_1M_1\lambda$.

Considering now the first and third factors, we may write down s_1s_3 linear constraints. Suppose they are again arrayed in numerical order, so that the constraint {$M_2\lambda$ equals the sum of all w_j for which j has u and v as its first and third components} has ordinal number s_3u+v+1. By adding the first s_3 relationships we find that the number of occurrences of level 0 of the first factor is $s_3M_2\lambda$. Similarly, by considering each of the s_1 sets of s_3 consecutive relationships, we find that the number of occurrences of any level of the first factor is $s_3M_2\lambda$. Therefore, M_2 equals $(s_2/s_3)M_1$, and what is more important, there are obvious linear dependencies between relationships written down by considering the first and second factors and those written down by considering the first and third factors. If every s_3-th relationship is removed or

"crossed out" from the set of $s_1 s_3$ relationships formed by considering the first and third factors, then these obvious dependencies will be removed.

Turning attention next to the second and third factors, one can write down $s_2 s_3$ relationships. The first, second, $\cdots$, $s_2 s_3$-th constraints are that $M_3 \lambda$ is equal to the sum of those w_j for which the second and third components of j are, respectively, 0 and 0, 0 and 1, $\cdots$, s_2 and s_3. Again, by adding sets of relationships, expressions can be obtained for the number of occurrences of each of the levels of the second factor and similarly for the third factor. But expressions are already available for each of these numbers of occurrences; those for each level of the second factor are obtainable from the first group of $s_1 s_2$ relationships, and those for the number of occurrences of each level of the third factor are obtainable from the second group of $s_1 s_3$ relationships. Suppose now that every s_3-th relationship is crossed out from the set of $s_2 s_3$ relationships formed by considering the second and third factors, and also the last s_3 relationships are crossed out. Then in all $s_2 + s_3 - 1$ constraints, and also all obvious dependencies, will be removed.

Thus far only three factors have been considered, but the above results can easily be extended to the remaining factors. From the set

of $s_1 s_i$ linear relationships obtained by considering the first and i-th factors, every s_i-th relationship may be crossed out in order to eliminate the obvious non-uniqueness of expressions for the number of occurrences of each level of the first factor. From the set of $s_i s_j$ relationships obtained by considering the i-th and j-th factors $(1 < i < j)$, every s_j-th relationship and the last s_j relationships may be crossed out, since expressions are already available for the number of occurrences of each level of both the i-th and j-th factors.

Before systematizing the crossing-out method for $t \geq 2$, we will show that the number of constraints remaining after using the crossing-out procedure is the same as that obtained from the base set in the case $t = 2$. There are $s_1 s_2$ relationships remaining in the first group, none having been removed. In the group obtained by considering the first and i-th factors there are $s_1 s_i - s_1$ constraints remaining. In the group obtained by considering the i-th and j-th factors there are $s_i s_j - s_i - s_j + 1$ constraints remaining. The total may be expressed as $s_1 s_2 + \Sigma_{i=3}^{n} s_1(s_i-1) + \Sigma_{i=2}^{n} \Sigma_{j=i+1}^{n} (s_i-1)(s_j-1)$. We have seen that the number of constraints using the base-set method is

$$1 + \Sigma_{i=1}^{n} (s_i-1) + \Sigma_{i=1}^{n} \Sigma_{j=i+1}^{n} (s_i-1)(s_j-1)$$
$$= 1 + \Sigma_{i=1}^{n} (s_i-1) + \Sigma_{i=2}^{n} \Sigma_{j=i+1}^{n} (s_i-1)(s_j-1) + (s_1-1) \Sigma_{i=2}^{n} (s_i-1)$$
$$= 1 + s_1 - 1 + s_1 \Sigma_{i=2}^{n} (s_i-1) + \Sigma \Sigma (s_i-1)(s_j-1)$$
$$= s_1 s_2 + s_1 \Sigma_{i=3}^{n} (s_i-1) + \Sigma \Sigma (s_i-1)(s_j-1).$$

Therefore, the methods yield equivalent sets of linearly independent constraints.

Formalization of the crossing-out method

We first formalize the procedure just described for $t = 2$. Consider each pair of factors in the standard (alphabetic) order BA, CA, CB, DA, DB, DC, EA, $\cdots$, where A, B, C, $\cdots$ are the first, second, third, $\cdots$ factors. We will say that a pair of factors is in class Ω if exactly Ω of the factors have occurred at least once in a previous pair. Thus, the pair BA is the only class 0 pair, the remaining pairs containing A are class 1 pairs, and the rest are class 2 pairs. Given the i-th and j-th factors, consider the constraints in the order obtained by looking at the pairs of levels in increasing numerical order 00, 01, $\cdots$, $0s_j$, 10, $\cdots$, $1s_j$, $\cdots$, s_i0, $\cdots$, s_is_j. Then if the pair is in class 0, cross out no constraints; if the pair is in class 1 cross out every s_j-th constraint; if the pair is in class 2 cross out every s_j-th constraint and the last s_j constraints.

Now suppose t is equal to 3. A triple of factors is in class Ω if Ω of the 3 pairs have occurred in a previous triple. If the triple consisting of the i-th, j-th, and k-th factors is in class Ω and if $\Omega \geq 1$, cross out every s_k-th constraint. If $\Omega \geq 2$, cross

out every s_j-th group of s_k constraints. If $\Omega = 3$, cross out the last $s_j s_k$ constraints.

In general, a t-tuple of factors is in class Ω if Ω of the (t-1)-tuples have occurred in previous t-tuples when written in standard alphabetic order. If the t-tuple consisting of the i_1-th through i_t-th factors is in class Ω and if $\Omega \geq \omega$, then cross out the last $\Pi_{j=t-\omega+2}^{t} s_{i_j}$ relationships in each set of $\Pi_{j=t-\omega+1}^{t} s_{i_j}$ consecutive relationships.

There are $\binom{n}{t}$ t-tuples, and the number of these in class Ω is $\binom{n-t-1+\Omega}{\Omega}$. The factors in a t-tuple in class Ω are the first through $(t-\Omega)$-th together with Ω of the factors numbered $t-\Omega+2$ through n. The number of constraints left after crossing out in a set of constraints formed by considering the t-tuple consisting of factors whose ordinal numbers are $1, \cdots, t-\Omega, i_1, \cdots, i_\Omega$, is $\{\Pi_{i=1}^{t-\Omega} s_i\} \times \{\Pi_{j=1}^{\Omega} (s_{i_j} - 1)\}$. The total number of constraints for $t = 2$ is given by $s_1 s_2 + \Sigma_{i_1=3}^{n} s_1(s_{i_1} - 1) + \Sigma_{i_1=2}^{n} \Sigma_{i_2=i_1+1}^{n} (s_{i_1} - 1)(s_{i_2} - 1)$, and for $t = 3$ by

$$s_1 s_2 s_3 + \Sigma_{i_1=4}^{n} s_1 s_2 (s_{i_1} - 1) + \Sigma_{i_1=3}^{n} \Sigma_{i_2=i_1+1}^{n} s_1 (s_{i_1} - 1)(s_{i_2} - 1)$$
$$+ \Sigma_{i_1=2}^{n} \Sigma_{i_2=i_1+1}^{n} \Sigma_{i_3=i_2+1}^{n} (s_{i_1} - 1)(s_{i_2} - 1)(s_{i_3} - 1).$$

The general expression can be written in the form

$\Pi_{i=1}^{t} s_i + \Sigma_{k=1}^{t}(\Pi_{i=1}^{t-k} s_i)\{\Sigma_{\theta'_{k,t}} \Pi_{j=1}^{k}(s_{i_j}-1)\}$, where the set $\theta'_{k,t}$ is the set of indices $i_1, \cdots, i_k$ such that $t-k+1 < i_1 < i_2 < \cdots < i_k \leq n$. The general expression derived above for the base-set method is

$1 + \Sigma_{k=1}^{t} \Sigma_{\theta_k} \Pi_{j=1}^{k} (s_{i_j}-1)$, where $\theta_k = \{i_j | 1 < i_1 < i_2 < \cdots < i_k \leq n\}$. These two expressions will now be shown to be equivalent.

THEOREM 7. For all t, the number of constraints using the base-set method is equal to the number of constraints using the crossing-out method.

PROOF. For fixed but arbitrary n and $s_1, s_2, \cdots, s_n$, let $g(t)$ be the number of constraints using the base-set method for a design of strength t and let $h(t)$ be the number of constraints using the crossing-out method. We have already shown that $g(2)$ equals $h(2)$. We will show that for all t, $g(t) - g(t-1)$ equals $h(t) - h(t-1)$, which, by induction, will complete the proof. The difference $g(t) - g(t-1)$ is given by $\Sigma_{\theta_t} \Pi_{j=1}^{t}(s_{i_j}-1)$, where $\theta_t = \{i_j | 1 \leq i_1 < i_2 < \cdots < i_t \leq n\}$. The difference $h(t) - h(t-1)$ is given by

$\Pi_{i=1}^{t} s_i + \Sigma_{k=1}^{t}(\Pi_{i=1}^{t-k} s_i)\{\Sigma_{\theta'_{k,t}} \Pi_{j=1}^{k} (s_{i_j}-1)\} - \Pi_{i=1}^{t-1} s_i$

$- \Sigma_{k=1}^{t-1} (\Pi_{i=1}^{t-k-1} s_i)\{\Sigma_{\theta'_{k,t-1}} \Pi_{j=1}^{k} (s_{i_j}-1)\}$. The latter may be rewritten

$h(t) - h(t-1) = \Sigma_{k=0}^{t-1} D_k + \{\Sigma_{\theta'_{t,t}} \Pi_{j=1}^{t} (s_{i_j} - 1)\}$, where
$D_k = (\Pi_{i=1}^{t-k} s_i)\{\Sigma_{\theta'_{k,t}} \Pi_{j=1}^{k} (s_{i_j} - 1)\} - (\Pi_{i=1}^{t-j_{k-1}} s_i)\{\Sigma_{\theta'_{k,t-1}} \Pi_{j=1}^{k}(s_{i_j} - 1)\}$.
The term D_0 reduces to $\Pi_{i=1}^{t-1} s_i(s_t - 1)$. When this is added to the first term of D_1 the result is
$(\Pi_{i=1}^{t-1} s_i)\{\Sigma_{i_1=t}^{n}(s_{i_1} - 1)\} = (\Pi_{i=1}^{t-1} s_i)\{\Sigma_{\theta'_{1,t-1}} \Pi_{j=1}^{1}(s_{i_j} - 1)$. Subtracting the second term of D_1 we have
$D_0 + D_1 = (\Pi_{i=1}^{t-2} s_i)(s_{t-1} - 1)\{\Sigma_{\theta'_{1,t-1}} \Pi_{j=1}^{1}(s_{i_j} - 1)\}$. Adding the first term of D_2 we obtain $(\Pi_{i=1}^{t-2} s_i)\{\Sigma_{\theta'_{2,t-1}} \Pi_{j=1}^{2}(s_{i_j} - 1)\}$, from which is derived $\Sigma_{k=0}^{2} D_k = (\Pi_{i=1}^{t-3} s_i)(s_{t-2} - 1)\{\Sigma_{\theta'_{2,t-1}} \Pi_{j=1}^{2}(s_{i_j} - 1)\}$. Proceeding analogously, the summation of the D_k becomes
$\Sigma_{k=0}^{t-1} D_k = (\Pi_{i=1}^{0} s_i)(s_1 - 1)\{\Sigma_{\theta'_{t-1,t-1}} \Pi_{j=1}^{t-1}(s_{i_j} - 1)\}$. We will now add this to the final term of $h(t) - h(t-1)$, namely $\{\Sigma_{\theta'_{t,t}} \Pi_{j=1}^{t}(s_{i_j} - 1)\}$. The index sets are $\theta'_{t-1,t-1} = \{1 < i_1 < i_2 < \cdots < i_{t-1} \leq n\}$ and $\theta'_{t,t} = \{1 < i_1 < i_2 < \cdots < i_t \leq n\}$. The result of the addition is therefore $\Sigma_{\theta_t} \Pi_{j=1}^{t}(s_{i_j} - 1)$, where θ_t is the set $\{1 \leq i_1 < i_2 < \cdots < i_t \leq n\}$, which is equal to $g(t) - g(t-1)$, as was to be proved.

Therefore, the base-set method and the crossing-out method are equivalent. Although the base-set method is used for the example in Section 3, the methods are equally easy to apply, and either may be used to construct a set of linearly independent constraints for the application of linear programming to the derivation of designs.

REFERENCES

ADDELMAN, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. Technometrics 4, 21-46.

ADDELMAN, S. and KEMPTHORNE, O. (1961). Orthogonal main-effect plans. ARL 9, Aeronautical Research Laboratory, Office of Aerospace Research, United States Air Force, Wright-Patterson Air Force Base, Ohio.

BOSE, R. C. and BUSH, K. A. (1952). Orthogonal arrays of strength two and three. Ann. Math. Statist. 23, 508-524.

BOX, G. E. P. and Hunter, J. S. (1961). The 2^{k-p} fractional factorial designs. Technometrics 3, 311-351 and 449-458.

BOX, G. E. P. and WILSON, K. B. (1951). On the experimental attainment of optimum conditions. J. R. Statist. Soc. B 13, 1-45.

CONNOR, W. S. and ZELEN, M. (1959). Fractional factorial experiment designs for factors at three levels. Nat. Bur. Standards Appl. Math. Ser. 54.

DANIEL, C. (1957). Fractional replication in industrial experimentation. Transactions of the Eleventh Annual Convention of the American Society for Quality Control, 229-233.

DYKSTRA, O. (1959). Partial duplication of factorial experiments. Technometrics 1, 63-75.

HADLEY, G. F. (1962). Linear Programming. Addison-Wesley, Reading, Mass.

PLACKETT, R. L. (1946). Some generalizations in the multifactorial design. Biometrika 33, 328-332.

PLACKETT, R. L. and BURMAN, J. P. (1946). The design of optimum multifactorial experiments. Biometrika 33, 305-325.

STEVENS, W. L. (1948). Statistical analysis of a non-orthogonal trifactorial experiment. Biometrika 35, 346-367.

WEBB, S. (1964). Optimality properties of orthogonal designs. Presented at the 124th Annual Meeting of the American Statistical Association, Dec. 27-30, 1964.

YATES, F. (1935). Complex experiments. J. R. Statist. Soc. (Suppl.) 2, 181-223.

APPENDIX B

CHARACTERIZATION OF NON-ORTHOGONAL INCOMPLETE FACTORIAL DESIGNS

SUMMARY

Two general classes of non-orthogonal incomplete factorial designs called clumpwise-orthogonal designs and permutation-invariant designs are defined. The cross-product matrices of the former can be arranged to contain blocks of non-zero elements down the main diagonal and zeros elsewhere. The latter class may be described as containing designs in which those factors which appear at the same number of levels are treated alike.

A review of the existing literature on non-orthogonal designs shows that despite the large quantity and variety in methods for construction, the designs all fall into one or both of the classes defined above.

It is shown that both classes of designs possess characterizations as integer solutions to sets of linear constraints. For a subclass called special clumpwise-orthogonal designs, defined only when all factors are at two levels, a different characterization is derived. This characterization involves group-theoretic considerations similar to those used in the classical theory of fractional factorials.

1. INTRODUCTION

Orthogonal incomplete factorial designs have been used for many years. In many experimental situations, however, the smallest orthogonal

incomplete design may entail more runs than the experimenter is prepared to make. Recently there has been increased interest in non-orthogonal designs, on which there is no a priori restriction on the number of runs, and many methods have been set forth for constructing such designs.

Criteria for comparing designs

Unlike orthogonal designs, which as a class enjoy certain optimality properties, non-orthogonal designs must be judged individually. First of all, the design must have at least the required resolution, which is defined as follows (Box & Hunter, 1961): If t is even and a design is of resolution $t+1$ then all parameters involving $\frac{1}{2}t$ or fewer factors are estimable, ignoring parameters involving more factors. If t is odd, parameters involving $(t-1)/2$ factors are estimable, ignoring parameters involving $(t+3)/2$ or more factors. Those involving $(t+1)/2$ factors are neither estimable nor completely ignored.

We will suppose that a particular parameterization has been decided upon. (See Webb (1963) for a discussion of parameterizations for factorial designs.) Given two designs which have the same resolution and number of runs, the one which permits estimation with smaller variances will, generally speaking, be more desirable. More specifically, four possible criteria for optimality of a design are:

i) that all the variances are minimized,

ii) that the trace of the covariance matrix is minimized,

iii) that the determinant of the covariance matrix is minimized, or

iv) that the average variance of a predicted value is minimized.

The first criterion is preferable, but often it is impossible to satisfy it with a design utilizing a given number of runs and a parameterization specified a priori. The second is equivalent to minimizing the sum of the variances of the estimates, and the third is equivalent to minimizing the volume of a confidence ellipsoid on the parameters (Mood, 1946). If the parameterization is scaled in an appropriate way, criterion (iv) is equivalent to (ii); more generally (iv) will reduce to a weighted average of the variances of the estimates.

Plackett and Burman (1946) showed that, subject to the restriction that the lengths of the columns of the design matrix be fixed, orthogonal designs satisfy criterion (i), hence also (ii), and (iv). Without this restriction, however, it is often possible, depending on the parameterization, to construct non-orthogonal designs which surpass orthogonal ones with the same number of runs under any of the last three criteria. In another paper (Webb, 1964), it was shown that, if one uses a main-effect parameterization under which the full factorial is orthogonal, criteria (iii) and (iv) are satisfied if and only if the design is also orthogonal.

Whereas knowledge of the conditions under which orthogonal designs are optimum is as yet incomplete, knowledge of such conditions for non-orthogonal designs is for the most part nonexistent. The few isolated cases in which any optimality property has been demonstrated for a non-orthogonal design are indicated in subsequent sections.

Classification of non-orthogonal designs

The non-orthogonal designs in the literature fall into two overlapping classes which I call permutation-invariant and clumpwise-orthogonal designs. A permutation-invariant design involving factors at the same number s of levels is one for which if abc...n is any treatment combination appearing in the design, where the symbols a, b, ..., n take on one of the values 0, 1, ..., s-1, then all the treatment combinations which are permutations of the given treatment combination also appear in the design. For the case in which all factors do not appear at the same number of levels, a permutation-invariant design is one which contains the same treatment combinations if factors appearing at the same number of levels are permuted. Later we will consider permutation-invariant designs of strength t, which are more general designs for which the cross-product matrix is not altered if factors appearing at the same numbers of levels are permuted. The parameter t is twice the number of factors in the highest-order interaction to be estimated.

A clumpwise-orthogonal design is one for which, by rearranging the columns of the coefficient matrix X, the cross-product matrix can be arranged so that there are square submatrices of non-zero elements down the main diagonal and zeros elsewhere. A subclass of particular interest, defined for the case when all factors appear at two levels, consists of the special clumpwise-orthogonal designs, which have all non-zero off-diagonal elements equal. I will restrict the definition of clumpwise-orthogonal designs by requiring that there be at least two clumps, since otherwise every design would satisfy the definition. It is convenient, however, not to impose this restriction on the class of special clumpwise-orthogonal designs.

Both the permutation-invariant and clumpwise-orthogonal classes contain orthogonal designs. A simple example of a design which is not orthogonal but which is still contained in both classes is the design for two-level factors consisting of runs at treatment combinations 00, 00, 01, 10, 11, and 11.

2. SURVEY OF NON-ORTHOGONAL DESIGNS

It is felt that this literature survey is reasonably exhaustive with the exception of quite recent work in the area. The non-orthogonal designs covered have been derived using many different devices and from widely different points of view. Still, they all may be categorized into the permutation-invariant and clumpwise-orthogonal classes.

Early non-orthogonal designs

Perhaps the earliest non-orthogonal designs were the optimum weighing designs proposed by Mood (1946). Suppose p objects are to be weighed on a single-pan scale. If each object is weighed individually the variance of each estimated weight is the error variance associated with each weighing. By weighing the objects in appropriately chosen groups, the variance of the estimated weights may be decreased. Using criterion (iii) (the determinant) to judge optimality, Mood showed that if $p = 2k-1$, then the design consisting of all combinations of the p objects taken k at a time is optimum. If $p = 2k$, he showed that the design which consists of all combinations taken k at a time together with all those taken $k + 1$ at a time is optimum. Banerjee (1948) pointed out that optimality is preserved if only those combinations which comprise a balanced incomplete block design are weighed. The designs of Mood and Banerjee are both permutation-invariant and special clumpwise-orthogonal with one clump.

Chakravarti (1956) introduced the "partially balanced array of strength t" which is a specialization of the permutation-invariant design of strength t to the case where all factors are at the same number of levels. I have chosen not to adopt his nomenclature because I believe "permutation-invariant" is a better description of the defining property of these designs. He gives two simple examples of such designs,

one containing runs at treatment combinations 0000, 0111, 1011, 1101, and 1110 for a 2^4 experiment, and another for a 2^5, containing runs at 00000 and all treatment combinations with four factors at their high level.

Morrison (1956) presented an interesting technique for constructing incomplete factorial designs in the case where all factors are not at the same number of levels. It is most effective when all but one of them are at two levels. The levels for the many-level factor are grouped into pairs (with one level left over if the number of levels is odd). A standard fractional factorial is constructed using, in turn, each pair with all of the two-level factors. To illustrate, consider an incomplete $2^4 5^1$ design, and designate the levels of the five-level factor by the integers 0 through 4. The design may be constructed in three stages. First, identifying levels 0 and 1 of the five-level factor with the levels of an (imaginary) fifth two-level factor E, an eighth-replicate of a 2^5 is constructed using as generators in the defining contrast AB, CD, and ACE. (For definitions see Daniel (1956).) Next, levels 2 and 3 are associated with E and another eighth-replicate is constructed using -AB, -CD, and ACE to generate the defining contrast. Finally, an eighth-replicate of the 2^4 is constructed, using AB, AC, and AD to generate the defining contrast, and level 4 of the five-level factor is used for these runs. The complete design is as follows:

0	0	0	0	0
1	1	1	1	0
0	0	1	1	1
1	1	0	0	1
0	1	0	1	2
1	0	1	0	2
0	1	1	0	3
1	0	0	1	3
0	0	0	0	4
1	1	1	1	4

Morrison's method of construction leads to clumpwise-orthogonal designs.

Irregular fractional factorials

A type of non-orthogonal design which has received a great deal of attention is the irregular fractional factorial. The general form of such an incomplete factorial is the k/s^m fraction of the s^n design. Banerjee (1949) discussed $(2^m-1)/2^m$ fractions of 2^n designs in connection with the weighing problem. Further examples were given by Kempthorne (1952) and by John (1961). Addelman (1961) gave a rather detailed analysis of such designs and presented a catalogue of $3/2^m$ replicates of 2^n designs. Dykstra (1959) gave a catalogue of irregular fractional factorials which involve partial duplication. All irregular fractional factorials are clumpwise-orthogonal, and those for the 2^n case are special clumpwise-orthogonal.

Recent methods for 2^m3^n experiments

Several workers have developed techniques for deriving incomplete 3^n and 2^m3^n designs, all of which lead to designs which are both

permutation-invariant and clumpwise-orthogonal. DeBaun (1959) developed a series of response-surface designs for three independent variables which are restricted to three levels, so that the designs are also incomplete factorials. The method is to combine the following subdesigns: the center point (treatment combination 111), the octahedron (011, 211, 101, 121, 110, and 112), the "cuboctahedron" (all treatment combinations with one 1) and the cube (all treatment combinations with no 1). DeBaun gave analyses for designs consisting of various combination of these subdesigns with repetition of entire subdesigns permitted, and he compared their efficiencies and variance contours as second-order response-surface designs.

Box and Behnken (1960), who, like DeBaun, were motivated by response-surface considerations, constructed a series of incomplete factorials utilizing balanced incomplete block configurations. The factors are associated with the block elements (often called "varieties" in the literature on balanced incomplete blocks). For each block a subdesign is constructed containing a two-level design using levels 0 and 2 of the factors appearing in the block, with all factors not in the block being held at level 1. After the procedure is repeated for every block the complete design is formed by combining the subdesigns and then appending several runs at the center point.

Connor (1960) developed a technique for handling 2^m3^n designs in which several combinations of standard fractions of the 2^m and of the 3^n are combined in a specified manner. One of his examples is the following incomplete $2^3 3^2$ design. Let the two half-replicates of the 2^3 be denoted symbolically by S_1 and S_2 and the three third-replicates of the 3^2 be denoted by T_1, T_2, and T_3. Let S_iT_j be the design formed by taking all possible combinations of treatment combinations in S_i with those in T_j. Then the design $S_1T_1 + S_2T_2 + S_2T_3$ is a half of the full factorial.

Another method for forming incomplete factorial designs for 3^n and 2^m3^n experiments was given by Fry (1961). The 3^n full factorial consists of a center point (the treatment combination 11...1) surrounded by concentric hyperspheres, the r-th containing the $\binom{n}{r}2^r$ treatment combinations with exactly r non-1 components. A design may be constructed by combining fractions of the treatment combinations on each hypersphere. Fry gives an example of an incomplete 3^4 design formed by taking all the points in every other hypersphere and omitting entirely the others (this is essentially DeBaun's method). Since the $\binom{n}{r}2^r$ points in the r-th hypersphere may be considered formally as $\binom{n}{r}$ separate factorials in r factors at two levels, standard fractions may be taken from each. Also, if the design contains m factors at two levels as well as the n three-level factors, fractions of the 2^{r+m}

design may be used. Fry gives as an example a "half-replicate" of a $2^3 3^2$ design constructed in this way. He points out that this method of construction leads to designs which are quite similar to those derived by Connor.

It is interesting to note that DeBaun, Box and Behnken, Connor, and Fry used disparate methods and were motivated by contrasting considerations. Yet the designs obtained are all permutation-invariant and clumpwise-orthogonal, and moreover usually possess a similar clump structure.

Expansible contractible designs

In work which has not previously been reported, K. W. Last[1] developed a series of permutation-invariant designs for factors at two levels. He was motivated by the need for designing experiments which may be prematurely halted, due for example to a breakdown of the equipment. His aim was to design the experiment to study first what are considered the most important factors so that if the design is not completed inferences may be made about the more interesting factors conditionally on the less important factors being fixed. A design is specified as the set of treatment combinations with a, b, ..., or k factors at the high level. Suppose the factors are arranged in increasing order of importance and the treatment combinations are run in increasing numerical

[1] Formerly with Rocketdyne, now with Sylvania Electric Products, Inc., Mountain View, California

order considered as binary numbers. Then the design will have the desired property. For example, consider three factors at two levels and let the design be the treatment combinations with zero, one, or three of the factors at their high level. If the combinations are run in the order 000, 001, 010, 100, 111, then after the first two observations the "main effect" (conditional on the other factors being fixed) of the last and most important factor can be estimated. After the third the conditional main effects of the latter two factors can be estimated, and finally after the fourth all three main effects can be estimated.

I proposed some specific permutation-invariant designs which are included in the general class considered by Last (Webb, 1961). These designs were (1) the incomplete 2^6 consisting of one run at each treatment combination containing zero, two, or five factors at the high level; (2) the incomplete 2^7 consisting of one run at each treatment combination containing zero, two, or six factors at the high level; and (3) the incomplete 2^7 consisting of one run at each treatment combination containing zero or two factors at the high level and two runs at each treatment combination with six factors at the high level.

Proportional-frequency designs

The general class of proportional-frequency designs has been investigated by Addelman (1962). Let n_{ij} be the number of occurrences of

the i-th level of one factor with the j-th level of another, $n_{i.}$ be the number of occurrences of the i-th level of the first factor, $n_{.j}$ be the number of occurrences of the j-th level of the second, and N be the number of runs. A design is a proportional-frequency design if, for every pair of factors and all levels i and j, $n_{ij}N$ equals $n_{i.}\ n_{.j}$.

If the grand mean is defined as the expected value of the average of the observations made, then proportional-frequency designs are orthogonal for estimating main effects and the grand mean. If, however, the grand mean is defined as the average of the expected values of all treatment combinations of the full factorial, then proportional-frequency designs are non-orthogonal. While neither the class of proportional-frequency designs nor the class of permutation-invariant designs includes the other, the designs which Addelman catalogues when considered as non-orthogonal designs are all contained in both classes.

3. A CHARACTERIZATION OF SPECIAL CLUMPWISE-ORTHOGONAL DESIGNS

Both permutation-invariant and clumpwise-orthogonal designs may be characterized as integer solutions to linear programming problems. Methods for constructing the constraint matrices for the general classes are given in the fourth and fifth sections. In the case of special clumpwise-orthogonal designs, defined only for the 2^n experimental

situations, a characterization is possible which is very different from that for the general class of clumpwise-orthogonal designs. This characterization, involving groups of treatment combinations, is reminiscent of the classical theory of fractional factorial designs.

Outline of the characterization

In the case of 2^n experiments, if a resolution $t + 1$ design is orthogonal, then for every subset of t factors every combination of levels occurs λ times (Webb, 1963). As I will show in this section, if a resolution $t + 1$ design is special clumpwise-orthogonal, then for every subset of t factors there exist constants λ_1 and λ_2 such that every combination of levels occurs either λ_1 or λ_2 times, and those appearing λ_1 times form a group. The proof will proceed by showing that, for each subset of t factors of a special clumpwise-orthogonal design, either a group of treatment combinations or the complement of a group can be added enough times to make the resulting augmented design orthogonal. In order to pursue this method of proof three preliminary results must be established. First we must show that if a set of treatment combinations forms a group under an appropriately defined operation, then it is a clumpwise-orthogonal design. Next we will show that any special clumpwise-orthogonal design in t factors leads to the same clump pattern as a group. Finally it must be established that the non-zero off-diagonal entry of a special clumpwise-orthogonal design in t factors is such that the group with the same

clump pattern or its complement can be added an appropriate integral number of times with the result that the augmented design is orthogonal.

Groups of treatment combinations

Fisher (1942) showed that the set of treatment combinations in a 2^n experiment form a group. Any given treatment combination may be designated as the group identity, and as such it is denoted by the symbol (1). Fisher makes no commitment, but later authors have chosen to let that treatment combination in which every factor appears at its low level be the identity. It will be more convenient to break with tradition and designate as the identity that treatment combination with every factor at its high level.

To each factor let there be assigned a letter a, b, ..., and let every treatment combination other than the identity be denoted by the letters corresponding to the factors at their low level. The group operation is defined as follows: the resultant of two treatment combinations under the group operation is that treatment combination which contains all the letters from the two omitting the letters they have in common.

Every treatment combination except the identity divides the group into two parts, those treatment combinations having an even number of letters in common with the given treatment combination, and those having

an odd number in common. The "even half" contains the identity (1) and is itself a subgroup. The set which has an even number of letters in common with each of two given treatment combinations will again be a subgroup, which will also have an even number of letters in common with both the product of the two given treatment combinations and the identity. Fisher's general result is that the set of elements even for every element of a subgroup of order 2^p is a subgroup of order 2^{n-p}, where n is the total number of factors. These two subgroups are said to be orthogonal.

The set of all possible parameters can also be considered as forming a group. The grand mean, usually denoted I in this context, is the identity. The main effects of the factors are denoted by A, B, C, and so on. Interactions are denoted by the appropriate combination of two, three, or more letters which belong to the factors interacting. The product of two symbols under the group operation, called the generalized interaction of the two symbols, is that symbol which contains all the letters of the two omitting any they may have in common. The identity I is considered as containing no letters. Thus, the generalized interaction of main effects A and B is AB, which is a genuine interaction. The generalized interaction of interactions BD and BDE is the main effect E. The group of parameters is isomorphic to the group of treatment combinations, so that Fisher's result applies to subgroups of the parameter group as well.

If a design is constructed whose treatment combinations form a subgroup of order 2^p or coset of a subgroup of order 2^p, the design will be a fractional factorial, that is, one block of a confounded design. The parameters aliased with the grand mean I are just those parameters in the subgroup orthogonal to the group of treatment combinations used. The parameters aliased with any given parameter are the elements of a coset of the orthogonal subgroup, where the coset is formed using the given parameter. The entire group of 2^n parameters is divided into 2^p alias sets, each containing $2^q = 2^{n-p}$ parameters. There is an estimable linear combination of the 2^q parameters in an alias set, and due to the particular choice of the identity, this estimable linear combination is just the sum of the parameters.

The cross-product matrices of subgroups of treatment combinations

Consider a coefficient matrix whose rows correspond to a subgroup of treatment combinations and which has columns corresponding to every possible main effect and interaction. The cross-product matrix formed from this coefficient matrix, which will be called the complete cross-product matrix associated with the subgroup, will of course be singular unless the subgroup considered is in fact the group itself.

THEOREM 1. The complete cross-product matrix of a subgroup of treatment combinations of order 2^p is special clumpwise-orthogonal, and all non-zero entries are 2^p.

PROOF. Finney (1945) showed that a subgroup defines an aliased design whose alias structure is determined by the alias subgroup (this fact is the basis of the classical theory of fractional replication). The coefficient vectors for aliased parameters are all equal, and coefficient vectors for parameters which are not aliased are orthogonal. Therefore, the complete cross-product matrix consists of 2^{n-p} clumps, where n is the number of factors, and each clump has all elements equal to 2^p.

THEOREM 2. The complete cross-product matrix associated with all treatment combinations except those of a subgroup of order 2^p is special clumpwise-orthogonal, with diagonal entries $2^n - 2^p$ and non-zero off-diagonal entries equal to -2^p.

PROOF. Represent the coefficient matrix of the full factorial in the form $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$, where X_2 corresponds to the part of the design whose treatment combinations form a group and where X_1 corresponds to the remainder. The complete cross-product matrix is $X_1'X_1 + X_2'X_2$, and we know already that it is diagonal with diagonal entries equal to 2^n. The matrix $X_2'X_2$ is known to be special clumpwise-orthogonal from Theorem 1, so that $X_1'X_1$ must also be special clumpwise-orthogonal with diagonal entries $2^n - 2^p$ and non-zero off-diagonal entries equal to -2^p.

Structure of special clumpwise-orthogonal designs

Let X be the coefficient matrix of a special clumpwise-orthogonal design of resolution $t + 1$ in n factors. All the elements of the cross-product matrix $X'X$ are of course inner products of pairs of columns of X. Denote by x_I the column of X which is the coefficient vector for the grand mean, by x_A, x_B, x_C, ... the coefficient vectors for the main effects, by x_{AB}, x_{AC}, ..., x_{BC}, ... the coefficient vectors for the two-factor interactions, and so forth. Because of the nature of coefficient vectors for the 2^n experimental situation, the squared length of every column of X is N, the number of runs. Because of the requirement that the design be of the special clumpwise-orthogonal form, the inner product of a pair of distinct columns is either zero or a single specified non-zero number, say c. Since the components of coefficient vectors for interactions are products of components of the coefficient vectors for the main effects of the factors making up the interactions, many inner products must be equal. For example, the inner product (x_I, x_{AB}) must be equal to (x_A, x_B) and to (x_{CD}, x_{ABCD}); (x_A, x_{BC}) must be equal to (x_B, x_{AC}) and to (x_C, x_{AB}). In fact, the inner products of pairs of x's whose subscripts have the same generalized interaction will be equal. In the two examples above, the common generalized interaction are AB in the first case and ABC in the second. Such a generalized interaction cannot involve more than t letters since the cross-product matrix comes from a resolution $t + 1$ design.

Since all pairs of x's whose subscripts have the same generalized interaction have the same inner product, we will discuss clumpwise-orthogonal designs in terms of which generalized interactions give rise to sets of non-zero inner products. If a generalized interaction gives rise to non-zero inner products, some clumping is induced in the cross-product matrix. To find what effects or interactions are clumped with a given effect or interaction,[1] it is only necessary to multiply the generalized interaction by the letters corresponding to the given effect or interaction. For example, if the generalized interaction AB gives rise to non-zero inner products, by multiplying AB by A we find that x_B must appear in the same clump as x_A. Continuing with our example, if the cross-product matrix is that of a design of resolution 5 or more, then x_I and x_{AB} are in the same clump. If the cross-product matrix is that of a design of resolution 7 or more, then x_C and x_{ABC}, x_D and x_{ABD}, and so on would be clumped together in pairs. The reason for the requirement that the resolution be large enough is that only interactions involving $\frac{1}{2}t$ or fewer factors are represented in the cross-product matrix of a resolution $t + 1$ design, so that in the second case above x_{AB} would not appear if the design were not treated as a design of at least resolution 5.

---

1. Aliasing is a limiting case of clumping in which an off-diagonal element is equal in absolute value to the diagonal elements. It should be noted that in the intermediate case when the off-diagonal element is in absolute value less than the diagonal elements, the corresponding parameters are in no sense "partially aliased" or "partially confounded". Rather, they are separately or jointly estimable.

If more than one generalized interaction gives rise to inner products equal to α, then their product must also, if this product contains t or fewer letters. For example, if the generalized interaction A and B both give rise to inner products equal to α, then x_I, x_A, and x_B are in the same clump, so that (x_A, x_B) must also be equal to α, which implies that all inner products of vectors whose generalized interaction is AB will also equal α. These results are proved generally in the Appendix as Theorem 3, which states that if two generalized interactions give rise to inner products equal to α, and if their product contains t or fewer letters, then their product must also give rise to inner products equal to α. The following theorem is an immediate corollary of Theorem 3.

THEOREM 4. For every subset of t factors in a special clumpwise-orthogonal design of resolution $t + 1$, the generalized interactions giving rise to non-zero inner products, together with I, form a group.

It follows from Theorem 4 that for any subset of t factors the complete cross-product matrix has the same clump structure as the complete cross-product matrix of a group. If the order of the group of generalized interactions mentioned in Theorem 4 is 2^q, then the complete cross-product matrix consists of $2^p = 2^{t-q}$ clumps each containing 2^q parameters. The parameters in the clump with the grand mean are just those generalized interactions which give rise to non-zero inner products,

so that the parameters in this clump form a subgroup of parameters. The parameters in the other clumps are cosets of this subgroup. We know already that there is a subgroup of treatment combinations whose complete cross-product matrix has exactly the same clump structure as the given clumpwise-orthogonal design. To be specific, it is that treatment-combination subgroup which is orthogonal to the subgroup of parameters in the clump which contains the grand mean.

Value of non-zero inner products

The value α of the non-zero inner products is related to the clump structure according to the following theorem. The proof is given in the Appendix.

THEOREM 5. In a special clumpwise-orthogonal design of resolution $t + 1$, if for a subset of t factors the group of generalized interactions giving rise to non-zero inner products is of order 2^q, then α is an integral multiple of 2^p.

We are now ready to characterize special clumpwise-orthogonal designs of resolution $t + 1$ containing only t factors. The general case of n factors will be discussed after the theorem.

THEOREM 6. A resolution $(t + 1)$ design in t factors with t even is special clumpwise-orthogonal if and only if there exist constants

λ_1 and λ_2 such that every treatment combination in a group of treatment combinations appears λ_1 times and every other treatment combination appears λ_2 times.

PROOF. The cross-product matrix formed from a partitioned coefficient matrix $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ is given by $\left[X_1'X_1 + X_2'X_2\right]$. That the stated condition on numbers of treatment combinations implies clumpwise-orthogonality is obvious from Theorems 1 and 2. Suppose a design containing N runs is clumpwise-orthogonal, that the group of generalized interactions giving rise to non-zero inner products is of order 2^q, and that the number of clumps is $2^p = 2^{t-q}$. Denote the coefficient matrix of the given special clumpwise-orthogonal design by X_1, and denote by X_2 the coefficient matrix of that group of treatment combinations which has the same clump structure as the given design. The group must be of order 2^p. Let $\alpha = c \cdot 2^p$, where c is, by Theorem 5, a non-zero integer.

CASE I. $\alpha < 0$. Consider a new design consisting of all treatment combinations of the original design augmented with the group of treatment combinations replicated $-c$ times. The cross-product matrix of the augmented design is $X_1'X_1 - c \cdot X_2'X_2$, which equals $(N - c \cdot 2^p)$ times the identity matrix of appropriate size. The augmented design is orthogonal, hence has every treatment combination replicated the same number of times, namely $(N - c \cdot 2^p)/2^t$. Therefore the original design must have been of the specified form with λ_1 equal to $\frac{N - c \cdot 2^p}{2^t} + c$,

and λ_2 equal to $\frac{N - c \cdot 2^p}{2^t}$. The total number of treatment combinations in the original design is $2^p(\frac{N - c \cdot 2^p}{2^t} + c) + (2^t - 2^p)(\frac{N - c \cdot 2^p}{2^t})$, which, as required, equals N.

CASE II. $\alpha > 0$. If the original design is augmented by adding the complement of the group c times, then the resulting design is an orthogonal design with every treatment combination replicated $\frac{N + c(2^t - 2^p)}{2^t}$ times. Hence the original design was of the specified form with $\lambda_1 = \frac{N + c(2^t - 2^p)}{2^t}$ and $\lambda_2 = \frac{N + c(2^t - 2^p)}{2^t} - c$.

Extension to n variables

According to Theorem 6, a necessary condition that a resolution $t + 1$ design in n variables be special clumpwise-orthogonal is that every sub-design containing only t of the factors have the structure specified in Theorem 6. It is necessary to impose some additional condition to make the converse true.

As an example of a design for which the converse does not apply, consider the following coefficient matrix and corresponding cross-product matrix:

$$X = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (X'X) = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 4 & 0 \\ 0 & 4 & 8 & 4 \\ 0 & 0 & 4 & 8 \end{bmatrix} .$$

Denote the four parameters by I, A, B, and C as before. By considering the subset consisting of factors one and two, the parameters A and B are clumped. Similarly, consideration of the second and third factors shows that B and C are clumped. The sub-design consisting of the first and third factors, however, is fully orthogonal, so that A and C do not appear in the same clump. In this case the clump structures of the three possible sub-designs are not mutually consistent.

The general characterization now follows from Theorem 6.

THEOREM 7. Consider a resolution $t + 1$ design in n factors with t even. If

a) for every subset of t factors there exist constants λ_1 and λ_2 satisfying the condition of Theorem 6,

b) the clump structures of the sub-designs are mutually consistent, and

c) the structures of the sub-designs yield the same non-zero off-diagonal element,

then the design is special clumpwise-orthogonal.

4. THE GENERAL CLASS OF CLUMPWISE-ORTHOGONAL DESIGNS

While it is not as simple or attractive a characterization as that just discussed, general clumpwise-orthogonal designs may be characterized as integer solutions to sets of linear constraints. Given a set of constraints for a fully orthogonal design, removal of one or more constraints yields a set of constraints for clumpwise-orthogonal designs. Alternatively, given a clump structure, a set of constraints may be written down expressing the orthogonality remaining in the design.

Deletion of constraints

For discussion of the way in which the removal of constraints induces clumps in the cross-product matrix, it will be helpful to use a $2^2 3^2$ design as an illustrative example. As usual, let the levels of the two-level factors be designated by 0 and 1 and the levels of the three-level factors by 0, 1, and 2. Treatment combinations will be written with the two-level factors appearing first. The symbol [ijkℓ] will be used to designate the number of times treatment combination ijkℓ appears in a design; [ijkx] will denote the sum [ijk0] + [ijk1] + [ijk2], and similarly for an x substituted into other positions.

It has been previously shown (Webb, 1963), that a design is an orthogonal design of resolution t + 1 (or strength t) if and only if for every subset of t factors, every combination of levels appears the

same number of times. A linearly independent set of constraints for a strength 2 design, obtained by the "base-set method" (Webb, 1963), is given in Table I.

We wish to find out what clumping is induced in the cross-product matrix if the first constraint is deleted. Since the constraints in the table are linearly independent, none of the remaining constraints is affected. It can be verified that for all permissible ij combinations, the following constraints still hold: $[ixjx] = 6\lambda$, $[ixxj] = 6\lambda$, $[xijx] = 6\lambda$. $[xixj] = 6\lambda$, and $[xxij] = 4\lambda$. Therefore, for the subset consisting of the first two factors the condition for orthogonality no longer holds, but for any other subset it is unaffected. It follows by the theorem on orthogonality previously stated that the design is orthogonal with the exception that the main effects of the two-level factors are clumped. The effect on the clump structure of removal of other constraints or sets of constraints can be found in an analogous manner.

Table I

$[00xx] = 9\lambda$	$[x0x2] = 6\lambda$
$[0x0x] = 6\lambda$	$[xx00] = 4\lambda$
$[0x1x] = 6\lambda$	$[xx01] = 4\lambda$
$[0xx0] = 6\lambda$	$[xx02] = 4\lambda$
$[0xx1] = 6\lambda$	$[xx10] = 4\lambda$
$[0xx2] = 6\lambda$	$[xx11] = 4\lambda$
$[x00x] = 6\lambda$	$[xx12] = 4\lambda$
$[x01x] = 6\lambda$	$[xx20] = 4\lambda$
$[x0x0] = 6\lambda$	$[xx21] = 4\lambda$
$[x0x1] = 6\lambda$	$[xx22] = 4\lambda$

Constraints for orthogonal $2^2 3^2$ design.

Constraints reflecting remaining orthogonality

The practice of removing one or more specific constraints from those for a fully orthogonal design cannot yield constraints for all clumpwise-orthogonal designs, for it may be that constraints for a particular clump structure can only be obtained by removing a particular linear combination of constraints. For example, for the $2^2 3^2$ experiment considered previously, constraints for the design with only the main effect of the first factor clumped with the mean cannot be constructed in the former manner. The alternative approach is to start with a clump structure and write down constraints reflecting the orthogonality remaining in the design.

To illustrate the procedure, the constraints for a particular clump structure for $2^2 3^2$ designs will be derived. Denote the grand mean by I, the effects of the two-level factors by A and B, and the linear and quadratic effects of the three-level factors by C, C^2, D, and D^2. The clump structure to be considered is described by the following diagrammatic representation of the cross-product matrix:

$$\begin{bmatrix} I & & & & & \\ & A & X & & & \\ & X & B & & & \\ & & & C & X & \\ & & & X & D & \\ & & & & & C^2 \quad X \\ & & & & & X \quad D^2 \end{bmatrix}.$$

Here the name of a parameter (I, A, B, etc.) is used to denote the squared length of the associated coefficient vector in the coefficient matrix. The symbol X denotes a non-zero off-diagonal element.

For this clump structure, any subdesign consisting of a two-level and a three-level factor is completely orthogonal. The corresponding constraints are those requiring that, for every subset consisting of one factor at two and one at three levels, every possible combination of levels occurs the same number of times λ. The corresponding constraints, reduced to a linearly independent set, are the first fifteen constraints given in Table II. In addition, C is required to be orthogonal to D^2 and D to C^2. The components of the coefficient vector for C are -1, 0, and 1 when C is at levels 0, 1, and 2 and similarly the components of the coefficient vector for D^2 are 1, -2, and 1 when D is at levels 0, 1, and 2. Therefore, the inner product of the coefficient vectors for C and D^2 is -[xx00] +2[xx01] - [xx02] + [xx20] -2[xx21] + [xx22]. The final two constraints in Table II are obtained by setting equal to zero the inner product of the coefficient vectors for C and D^2 and likewise the inner product of the coefficient vectors for D and C^2.

Other clump structures for other designs can be handled in an analogous manner. This concludes the discussion of clumpwise-orthogonal designs.

Table II

$[0x0x] = \lambda$	$[x00x] = \lambda$	$[1x0x] = \lambda$
$[0x1x] = \lambda$	$[x01x] = \lambda$	$[1x1x] = \lambda$
$[0x2x] = \lambda$	$[x02x] = \lambda$	$[1x2x] = \lambda$
$[0xx0] = \lambda$	$[x0x0] = \lambda$	$[1xx0] = \lambda$
$[0xx1] = \lambda$	$[x0x1] = \lambda$	$[1xx1] = \lambda$

$$-[xx00] + 2[xx01] - [xx02] + [xx21] - 2[xx21] + [xx22] = 0$$

$$-[xx00] + [xx02] + 2[xx10] - 2[xx12] - [xx20] + [xx22] = 0$$

Constraints for $2^2 3^2$ clumpwise-orthogonal design.

5. THE CLASS OF PERMUTATION-INVARIANT DESIGNS

The construction of constraints for clumpwise-orthogonal designs involves weakening the constraints for orthogonal designs by deleting linear combinations of constraints. The constraints for permutation-invariant designs, on the other hand, are obtained by introducing more indices. Before showing how this is accomplished, some additional general material on this class of designs will be presented.

Permutation-invariant designs of strength t

The appealing property of permutation-invariant designs is that factors with the same number of levels are treated alike in the sense that the estimates of corresponding main effects have the same variances and the same covariances with other estimates. For this to hold in a design of resolution $t + 1$, it is not necessary that the design be permutation-invariant as a whole, but only that all subdesigns containing t factors be permutation-invariant (this follows from the fact that the cross-product matrix contains inner products of interactions involving up to $\frac{1}{2}t$ factors, so that the largest number of factors involved in computing any element of this matrix is t). Therefore, it is convenient to define formally a more general class of designs, permutation-invariant designs of strength t.

Given n factors, select any ordered subset of τ factors and let $f(i_1, i_2, \ldots, i_\tau)$ be the number of occurrences of level i_1 of the first factor of the subset with level i_2 of the second, ..., with level i_τ of the τ-th. Suppose t is the largest value of τ such that the function $f(i_1, i_2, \ldots, i_\tau)$ depends on the number of levels of the factors chosen to be first through τ-th, but not on the particular choice of ordered subset. Then the design is a permutation-invariant design of strength t. It should be noted that the permutation-invariant designs mentioned in the literature survey are all of full strength; that is, they are of strength n where n is the number of factors.

Several permutation-invariant designs can be shown to satisfy one or another of the criteria for optimality given in the introduction. It can be shown by an onumeration of all possible five-run incomplete 2^4 designs that the design given by Chakravarti (1956) consisting of runs at treatment combinations 0000, 0111, 1011, 1101, and 1110 is the essentially unique resolution 3 design which satisfies criterion (i) (the variances are each minimized). By "essentially unique" I mean that the only other designs with this property are those derived from the given design by interchanging for some of the factors the designation of which is the high and which is the low level. The complement of this design, that is, the design consisting of runs at the other eleven treatment combinations, is the essentially unique eleven-run resolution 5 design which

satisfies criteria (ii), (iii), and (iv) (the trace, determinant, and average predicted value are minimized). The twelve-run design obtained by adding a run at the treatment combination 0000 is still permutation-invariant and also satisfies criteria (ii), (iii), and (iv) among twelve-run designs (Webb, 1961). It is not unique in this respect, however, there being three other twelve-run designs, not essentially equivalent, which also satisfy these three criteria. Two of the other designs, both irregular fractional factorials and hence special clumpwise-orthogonal, are that obtained by omitting 0000, 0011, 1110, and 1101, and that obtained by omitting 0000, 0110, 0101, and 0011. The third design, neither clumpwise-orthogonal nor permutation-invariant, omits 0000, 1110, 1101, and 1011.

Linear constraints for permutation-invariant designs

For orthogonal designs the total number of runs in certain subsets of treatment combinations is equal to a known multiple of the single index λ. For permutation-invariant designs, the total number of runs at treatment combinations in these subsets is equal to a known multiple of one of several indices. To illustrate, constraints for a $2^3 3^2$ permutation-invariant design of strength 2 will be derived.

According to the definition, the number of times a pair of two-level factors appear at levels 0 and 0 is independent of the choice of the

pair. Denoting this number by α_1, we obtain the constraints [OOxxx] = α_1, [OxOxx] = α_1, and [xOOxx] = α_1. Similarly, the number of times a pair of two-level factors appear at levels 0 and 1 (or equivalently at 1 and 0) is a constant α_2 which may differ from α_1. The number of times a pair of two-level factors appear at levels 1 and 1 is α_3. By considering the three-level factors six more indices may be defined. Let $\beta_1, \ldots, \beta_6$ be the number of occurrences of the combinations 00, 01, 02, 11, 12, and 22 respectively. Finally, consideration of one factor at two levels and one at three levels produces the indices $\gamma_1, \ldots, \gamma_6$ corresponding to the combinations 00, 01, 02, 10, 11, and 12 respectively. A set of linear constraints may now be written down. As usual, the complete set is linearly dependent, and in Table III a linearly independent subset has been extracted.

As was noted previously, an orthogonal design is a special case of a permutation-invariant design. More specifically, an orthogonal design of strength t is a permutation-invariant design of strength t in which all the indices are the appropriate multiples of the single index λ. It may be verified that if in Table III α_1 is replaced by 9λ, each β by 4λ, and each γ by 6λ, the constraints are those of a strength 2 orthogonal design.

Table III

$[00xxx] = \alpha_1$	$[x0x1x] = \gamma_2$	$[xxx00] = \beta_1$
$[0x0xx] = \alpha_1$	$[x0xx0] = \gamma_1$	$[xxx01] = \beta_2$
$[x00xx] = \alpha_1$	$[x0xx1] = \gamma_2$	$[xxx02] = \beta_3$
$[0xx0x] = \gamma_1$	$[x0xx2] = \gamma_3$	$[xxx10] = \beta_2$
$[0xx1x] = \gamma_2$	$[xx00x] = \gamma_1$	$[xxx11] = \beta_4$
$[0xxx0] = \gamma_1$	$[xx01x] = \gamma_2$	$[xxx12] = \beta_5$
$[0xxx1] = \gamma_2$	$[xx0x0] = \gamma_1$	$[xxx20] = \beta_3$
$[0xxx2] = \gamma_3$	$[xx0x1] = \gamma_2$	$[xxx21] = \beta_5$
$[x0x0x] = \gamma_1$	$[xx0x2] = \gamma_3$	$[xxx22] = \beta_6$

Constraints for $2^3 3^2$ permutation-invariant design.

APPENDIX

This Appendix contains proofs of two Theorems used in Section 3.

Recall that the column of the coefficient matrix X which is the coefficient vector for the parameter j is denoted by x_j. Here j may be I (the grand mean), A, B, ..., AB, AC, ..., etc. Two pairs of columns have the same inner product if their subscripts have the same generalized interaction. Special clumpwise-orthogonal designs can therefore be characterized in terms of which generalized interactions give rise to inner products equal to the non-zero off-diagonal element α. For a design of resolution $t + 1$ only parameters involving $\frac{1}{2}t$ or fewer factors are involved in the cross-product matrix. The generalized interactions giving rise to non-zero inner products may contain up to t letters. If such a generalized interaction is multiplied by a given effect or interaction, and if the resultant contains $\frac{1}{2}t$ or fewer letters, then the resultant and the given parameter appear in the same clump.

THEOREM 3. If two generalized interactions give rise to inner products equal to α, and if their product contains t or fewer letters, then in order for the design to be special clumpwise-orthogonal their product must also give rise to inner products equal to α.

PROOF. Let c be the number of letters common to both generalized interactions, d be the number of unique letters in one, and e the

number of unique letters in the other. Since generalized interactions giving rise to non-zero inner products must contain t or fewer letters, we have $c + d \leq t$ and $c + e \leq t$. Also $d + e \leq t$, since by hypothesis the product of the two generalized interactions contains no more than t letters. Without losing generality it may be assumed that $d \leq \frac{1}{2} t$. The theorem can now be proved by exhibiting for each of several cases a pair of x's which must be in the same clump, and the generalized interaction of whose subscripts is the product of the two original generalized interactions. Let $v = \frac{1}{2} t$, so that only x's with v or fewer letters in their subscripts can be considered. Let the two generalized interactions be denoted by $(i_1, \ldots, i_c, j_1, \ldots, j_d)$ and $(i_1, \ldots, i_c, k_1, \ldots, k_e)$.

CASE I. $c > v$, d and e unrestricted. By multiplying the two generalized interactions by $(i_1, \ldots, i_v)$ it is found that the three x's with subscripts $(i_1, \ldots, i_v)$, $(i_{v+1}, \ldots, i_c, j_1, \ldots, j_d)$, and $(i_{v+1}, \ldots, i_c, k_1, \ldots, k_e)$ are all in the same clump. Hence the inner product of the last two must be equal to α, so that all pairs of x's whose subscripts have the generalized interaction $(j_1, \ldots, j_d, k_1, \ldots, k_e)$ must have inner product α.

CASE II. $c \leq v$, $e \leq v$, d unrestricted. By multiplying the two generalized interactions by $(i_1, \ldots, i_c)$ it is found that the three x's with subscripts $(i_1, \ldots, i_c)$, $(j_1, \ldots, j_d)$, and $(k_1, \ldots, k_e)$

must all be in the same clump. Therefore all pairs of x's whose subscripts have generalized interaction $(j_1, \ldots, j_d, k_1, \ldots, k_e)$ must have their inner product equal to α. If c is zero, $(i_1, \ldots, i_c)$ is taken to be the single subscript I.

CASE III. $c \leq v$, $e > v$, d unrestricted. By multiplying the two generalized interactions by $(i_1, \ldots, i_c, k_1, \ldots, k_{v-g})$, where g is the larger of c and d, it is found that the x's with subscripts $(i_1, \ldots, i_c, k_1, \ldots, k_{v-g})$, $(j_1, \ldots, j_d, k_1, \ldots, k_{v-g})$, and $(k_{v-g+1}, \ldots, k_e)$ are in the same clump. Note that the number of letters in these three generalized interactions are $v - g + c$, $v - g + d$, and $e - v + g$ and these three numbers are all less than or equal to v, so that the three x's named all appear in the X matrix.

THEOREM 5. In a special clumpwise-orthogonal design of resolution $t + 1$, if for a subset of t factors the group of generalized interactions giving rise to non-zero inner products is of order 2^q, then α is an integral multiple of 2^p, where $p = t - q$.

PROOF. Given a subset of t factors, suppose it is possible to pick two x's from the complete design matrix which are not in the same clump with each other or with x_I. Then the x whose subscript is the generalized interaction of the first two x's must be in a fourth clump, for if not, the first two x's would either be in the same clump

or one would be in the clump containing x_I. In general, then, it is possible to pick a set of 2^p x's, one from each clump, whose subscripts form a group. But this set of 2^p x's itself forms a design matrix of a resolution $2p + 1$ design in p factors which can be taken to be any set of p generators of the group. Because each x comes from a separate clump, it is an orthogonal design. By a result of Webb (1963), this design must consist of the full factorial replicated some number, k, of times. Therefore the coefficient matrix X of this subdesign consists of 2^p sets of k equal row vectors.

Now consider any vector x_i which is in the same clump with x_I. Since x_i is orthogonal to each column (except x_I) of the coefficient matrix for the orthogonal subdesign, the projection of x_i on the subspace spanned by the columns of the orthogonal subdesign must be a constant times x_I. This subspace consists of all vectors which are constant over each set of k consecutive components, so that the projection of x_i onto this subspace replaces each component in a set by the average of the k components in the set. Since the projection is a multiple of x_I, the sum of sets of k consecutive components of x_i must be a constant c, and since x_i consists of plus and minus ones, c must be integral. But c is equal to the inner product of x_I and x_i, which is the sum of all the components of x_i, which must be c times 2^p.

REFERENCES

ADDELMAN, S. (1961). Irregular fractions of the 2^n factorial experiments. Technometrics 3, 479-496

ADDELMAN, S. (1962). Orthogonal main-effect plans for asymmetrical factorial experiments. Technometrics 4, 21-46

BANERJEE, K. S. (1948). Weighing designs and balanced incomplete blocks. Ann. Math. Statist. 19, 394-399.

BANERJEE, K. S. (1949). A note on weighing designs. Ann. Math. Statist. 20, 300-304.

BOX, G. E. P. and BEHNKEN, D. W. (1960). Some new three level designs for the study of quantitative variables. Technometrics 2, 455-475.

BOX, G. E. P. and HUNTER, J. S. (1961). The 2^{k-p} fractional factorial designs. Technometrics 3, 311-351 and 449-458.

CHAKRAVARTI, I. M. (1956). Fractional replication in asymmetrical factorial designs and partially balanced arrays. Sankhyā 17, 143-146.

CONNOR, W. S. (1960). Construction of fractional factorial designs of the mixed 2^m3^n series, 168-181. Contributions to Probability and Statistics Ed. I. Olkin, et al. Stanford Univ. Press.

DANIEL, C. (1956). Fractional replication in industrial research. Proc. Third Berkeley Symp. Math. Statist. Probability 5, 87-98. Univ. of California Press, Berkeley.

DE BAUN, R. M. (1959). Response surface designs for three factors at three levels. Technometrics 1, 1-8.

DYKSTRA, O. (1959). Partial duplication of factorial experiments. Technometrics 1, 63-75.

FINNEY, D. J. (1945). The fractional replication of factorial arrangements. Ann. Eugenics 12, 291-301.

FISHER, R. A. (1942). The theory of confounding in factorial experiments in relation to the theory of groups. Ann. Eugenics 11, 341-353.

FRY, R. E. (1961). Finding new fractions of factorial experimental designs. Technometrics 3, 359-370.

JOHN, P. W. M. (1961). Three-quarter replicates of 2^4 and 2^5 designs. Biometrics 17, 319-321.

KEMPTHORNE, O. (1952). The Design and Analysis of Experiments. Wiley, New York

MOOD, ALEXANDER M. (1946). On Hotelling's weighing problem. Ann. Math. Statist. 17, 432-446.

MORRISON, M. (1956). Fractional replication for mixed series. Biometrics 12, 1-19.

WEBB, S. (1961). Some new incomplete factorial designs. Paper presented at the Urbana, Illinois meeting of the Inst. Math. Statist., November 24 and 25, 1961.

WEBB, S. (1963). Orthogonal incomplete factorial designs and linear programming. Research Report 63-5, Rocketdyne, Canoga Park.

WEBB, S. (1964). Optimality properties of orthogonal designs. Presented at the 124th annual meeting of the American Statistical Association, Dec. 27 - 30, 1964.

APPENDIX C

OPTIMALITY PROPERTIES OF ORTHOGONAL DESIGNS

SUMMARY

It is well known that, subject to the restriction that the squared lengths of the column vectors of the design matrix be fixed, the variances of the individual parameter estimates are each minimized if and only if an incomplete factorial design is orthogonal. Except in the case of 2^n experiments the restriction will often be artificial, and without the restriction it is shown that this particular optimality property no longer applies.

Two criteria for optimality will be considered: a) minimization of the volume, adjusted for the number of runs, of a confidence ellipsoid on the estimated parameters, and b) minimization of the average variance of a predicted value adjusted for the number of runs, where the average is taken over all points of the full factorial. Considering parameterizations for which the full factorial is orthogonal and which involve main effects only, it is shown that criteria (a) and (b) are both satisfied if and only if the design is orthogonal.

These criteria are also satisfied for many more complicated parameterizations, including main effects plus all two-factor interactions, main effects plus all two- and three-factor interactions, etc. They also are satisfied for any nonsingular transformation of any of these parameterizations.

1. INTRODUCTION

This paper is concerned with factorial designs, in which a number of factors are each restricted a priori to a specified number of levels, in contrast to regression designs, in which the factors are continuously variable. Whether or not a factorial design is orthogonal depends on the parameterization used to describe the response. The expected values of the response of interest will vary between the points or treatment combinations of the full factorial, and the parameters are defined in terms of these expected responses. A canonical parameterization will be specified in the following manner: The grand mean $\bar{\mu}$ is the average expected response, averaged over every point of the full factorial. The main effects for a factor at say s levels are defined in terms of the quantities $\bar{\mu}_i$, $i = 0, 1, \ldots, s - 1$, which are the expected responses averaged over all treatment combinations in which the given factor is at level i. More specifically, there are $s - 1$ main-effect parameters each defined as a linear combination of the $\bar{\mu}_i$, and such that the linear combinations are orthogonal and the sums of their coefficients are zero. It will be convenient for later use to let the linear combinations be scaled so that each has squared length $1/s$. Interaction parameters involving two or more factors are similarly defined in terms of average expected values with the factors involved at fixed levels.

Once a set of parameters is defined in terms of the expected responses at the various treatment combinations, the process can be reversed, and the expected responses may be expressed approximately as linear combinations of the parameters. If there are as many parameters as points in the full factorial (that is, if interactions of all orders are defined),

then the expected responses can be expressed exactly. If there are fewer parameters, then some degree of approximation is necessary. In practice, however, the approximation is often good, relative to experimental error, even when the number of parameters is much smaller than the number of points.

Given N observations at any set of treatment combinations, the values of the responses can be expressed in the form $Y = X\beta + e$. Here Y is the $N \times 1$ vector of responses, β is the $p \times 1$ vector of parameters, X is an $N \times p$ matrix called the coefficient matrix, and e is an $N \times 1$ vector of errors. If the approximation of expected values by linear combinations of parameters is adequate, then the components of e are assumed to be independent and identically distributed experimental errors. If the matrix $X'X$ is nonsingular, then the least-squares estimate of the vector β is $\hat{\beta} = (X'X)^{-1}X'Y$. The covariance matrix of $\hat{\beta}$ is $(X'X)^{-1}\sigma^2$, where σ^2 is the common unknown variance of the components of e.

2. CRITERIA FOR OPTIMALITY

Consider a class of designs for which the lengths of the column vectors of X are fixed. If there is an orthogonal design (i.e. $X'X$ is diagonal) in this class, then the variances of all the estimates will be smallest using this design. This is a well known result which was proved by Plackett and Burman[1]. The restriction that the lengths of the columns be fixed may appear to be simply a natural way of fixing the scale of measurement, but for factorial designs it is a very artificial restriction. If, as is more natural, the number of runs and the parameterization are fixed, then this optimality property of orthogonal designs no longer holds, as the example in the following paragraph indicates.

An experimental situation will be considered in which there is a single factor at three levels. This example is chosen for simplicity, but it is nontrivial in the sense that the conclusions which can be drawn from it apply equally well to more comprehensive experimental situations. The parameters, defined in terms of the expected responses μ_0, μ_1, and μ_2, are the grand mean $\beta_0 = (\mu_0 + \mu_1 + \mu_2)/3$, the "linear effect" $\beta_1 = (\mu_2 - \mu_0)/\sqrt{6}$, and the "quadratic effect" $\beta_2 = (\mu_0 - 2\mu_1 + \mu_2)/3\sqrt{2}$. The design containing two runs at each treatment combination is orthogonal and the variances are $\sigma^2/6$ for each of the three estimates. The design containing two runs at level 0, one run at 1, and three runs at 2 has variances $11\sigma^2/54$ for $\hat{\beta}_0$, $5\sigma^2/36$ for $\hat{\beta}_1$, and $29\sigma^2/108$ for $\hat{\beta}_2$. Thus the variance for $\hat{\beta}_1$ is smaller than for the orthogonal design with the same number of runs. The design with one run at level 0, three runs at 1, and two runs at 2 yields variances of $11\sigma^2/54$ for $\hat{\beta}_0$, $\sigma^2/4$ for $\hat{\beta}_1$, and $17\sigma^2/108$ for $\hat{\beta}_2$. If one uses this design, the variance of $\hat{\beta}_2$ is smaller than for the orthogonal design.

It is apparent from this example that the given criterion for optimality is too strong; that is, there will in general be no design for which the variance of each estimate will be minimized. It therefore seems appropriate to search for other criteria for optimality.

The purpose of a factorial experiment is usually either to estimate parameters or to describe the response over the grid of possible treatment combinations. Consequently, it seems natural to consider two criteria corresponding to these two situations.

Consider designs of N runs with p factors. The first criterion is based on the product of the determinant of the matrix $(X'X)^{-1}$ and N^p. If

an N-run design is replicated r times, so that the complete design contains Nr runs, each element of the covariance matrix $\sigma^2(X'X)^{-1}$ of the original design is divided by r and the determinant of $(X'X)^{-1}$ is divided by r^p. The purpose of multiplying the determinant by N^p is therefore to adjust for the number of runs. The reason for using the determinant in the first place is as follows. A confidence set for the parameters is given by the set of points for which $(\beta-\hat{\beta})'(X'X)(\beta-\hat{\beta}) \leq ps^2K$, where s^2 is the residual sum of squares and where K is a constant depending on the confidence level $(1-\alpha)$, p, N, and the distribution of the errors. If the errors are normally distributed, K is the upper α point of the F distribution with p and $N-p$ degrees of freedom. The volume of the confidence ellipsoid so determined is given by

$$V = \frac{2\pi^{\frac{1}{2}p} \quad (ps^2K)^{\frac{1}{2}p}}{p\Gamma(\frac{1}{2}p)\sqrt{\det(X'X)}} \quad .$$

Therefore, the volume of this ellipsoid is a monotonically increasing function of the determinant of $(X'X)^{-1}$.

The second criterion is appropriate when the purpose of the experiment is to predict the response equally well at each treatment combination of the full factorial. Such predictions can be made using the parameter estimates, and the variances of the predictions depend on the variances and covariances of these estimates. The criterion proposed is the average variance of a predicted value, times the number of runs. Multiplication by N again has the effect of scaling for the number of runs.

3. MAIN-EFFECT PARAMETERIZATIONS

In this section only main-effect parameterizations will be considered, but the number of factors and the numbers of levels of each factor are arbitrary.

Let Z be the coefficient matrix of the full factorial design consisting of p factors and let M be the number of runs in this design. Due to the particular choice of scaling used in defining the parameters, the following theorem is true.

THEOREM 1. The squared length of every column vector of Z is M and the squared length of every row vector is p.

PROOF. Consider an arbitrary factor at any arbitrary number s of levels. With this factor are associated $s - 1$ columns of Z, the coefficient vectors for the main effects of this factor. The components of these coefficient vectors may be determined from the definition of the main-effect parameters. Let β_0 represent the grand mean and $\beta_1, \ldots, \beta_{s-1}$ represent the main effects of the given factor. These parameters are defined as linear combinations of $\bar{\mu}_0, \ldots, \bar{\mu}_{s-1}$. Let $Q = \{q_{ij}\}$ be the matrix of the coefficients of these linear combinations, so that we have

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{s-1} \end{bmatrix} = \begin{bmatrix} 1/s & 1/s & \cdots & 1/s \\ q_{21} & q_{22} & \cdots & q_{2s} \\ \vdots & \vdots & & \vdots \\ q_{s1} & q_{s2} & \cdots & q_{ss} \end{bmatrix} \begin{bmatrix} \bar{\mu}_0 \\ \bar{\mu}_1 \\ \vdots \\ \bar{\mu}_{s-1} \end{bmatrix}$$

which may be written in matrix notation as $\beta^* = Q\mu^*$ (by making the obvious definitions). Note that the elements of the first row of the matrix Q are all $1/s$, and that this is consistent with the definition of the grand mean given in the introduction. By definition the rows of Q are orthogonal and have squared length $1/s$.

Consider the matrix $P = Q^{-1}$. It can be verified that due to the orthogonality of the columns of Q, P is as follows

$$P = \begin{bmatrix} 1 & sq_{21} & \cdots & sq_{s1} \\ 1 & sq_{22} & \cdots & sq_{s2} \\ \vdots & \vdots & & \vdots \\ 1 & sq_{2s} & \cdots & sq_{ss} \end{bmatrix}$$

or $P = sQ'$. Now $PP' = P'P = sI$, so that the sum of squares of any row or column of P is s.

Consider now the submatrix consisting of the columns of Z associated with the main effects of the factor. Any row of this submatrix has the same elements as a row of P with the first element omitted. Since the design is the full factorial, each level of the factor appears the same number of times, and hence each row of P is repeated the same number of times k. The constant k is the product of the number of levels of the remaining factors. The sum of squares of each column of the submatrix is therefore $k \times s$, which equals M. The sum of squares of each row of the submatrix is obviously $s - 1$.

Finally, consider the entire matrix Z. The squared length of columns associated with main effects has been shown to be M and the column associated with the grand mean has 1 for all its components, hence also has squared length M. It has also been shown that the squared length of the part of a row associated with the i-th factor is $s_i - 1$, so that the total squared length is $1 + \sum_{i=1}^{n} (s_i - 1)$, which equals p, the number of parameters. The proof is now complete.

Using Theorem 1 and the facts that the trace and the determinant of a matrix are, respectively, the sum and the product of its eigenvalues, we can prove the following theorem.

THEOREM 2. Both criteria [(a) and (b)] are satisfied if and only if a design is orthogonal under the canonical main-effect parameterization, in which case they will also be satisfied under any other main-effect parameterization, and the variances of all predicted values are equal.

PROOF. Consider first the average variance of a predicted value times the number of runs. This is given by (σ^2N/M) trace $Z(X'X)^{-1}Z' = \sigma^2N$ trace $(X'X)^{-1} = \sigma^2N \sum_{i=1}^{p} 1/\lambda_i$, where the λ_i are the eigenvalues of $X'X$. Since by Theorem 1 the sums of squares of elements of the rows of Z are all p, the same is true of X, and we have trace XX' = trace $X'X = Np$, so that $\Sigma\lambda_i$ = trace $X'X$ is fixed. In order to minimize $N\Sigma 1/\lambda_i$ with $\Sigma\lambda_i$ fixed, the λ_i must all be equal to N, which implies that $X'X$ is a scalar multiple of the identity matrix. In this case the variances of all predicted values (the diagonal elements of $\sigma^2Z(X'X)^{-1}Z'$) are equal to σ^2p/N, and N times the average is σ^2p/M. The chosen determinant criterion is $N^p\det(X'X)^{-1}$. To minimize $\det(X'X)^{-1}$ with $\Sigma\lambda_i$ fixed, the λ_i must again all be equal, and the value of the criterion is unity.

Now consider any nonsingular reparameterization $\alpha = A^{-1}\beta$. The expression for the first criterion takes the form (σ^2N/M) trace $ZA(A'X'XA)^{-1}A'Z'$, so that the A's cancel and the value is the same as before. The second criterion becomes $N^p\det(X'X)^{-1}/(\det A)^2$, so that under a new parameterization both criteria are minimized if and only if the design is orthogonal under the canonical parameterization. This completes the proof.

4. HIGHER-ORDER PARAMETERIZATIONS

In the proof of Theorem 2 the only reason for restricting the parameterization to contain only main effects is so that Theorem 1 will apply. Therefore, Theorem 2 applies to any canonical parameterization for which Theorem 1 holds. It will now be demonstrated that Theorem 1 is valid for many canonical parameterizations involving interactions.

Let Z be the coefficient matrix of the full factorial under a canonical parameterization including the grand mean, main effects, and two-factor interactions. Consider those columns corresponding to the main effects and the two-factor interactions involving factors at s_1 and s_2 levels. Treating these factors and their interactions as a single $s_1 s_2$-level factor, we find that Theorem 1 applies and that the squared lengths of these columns are all equal to M and that the squared length of the rows is $s_1 s_2 - 1$. We know from Theorem 1 that the portion of this squared length attributable to the columns for the main effects of the two factors is $(s_1-1) + (s_2-1)$. The remainder is $(s_1-1)(s_2-1)$, which is equal to the number of interaction parameters. This argument may be repeated for the interactions between any pair of factors, or for that matter for interactions involving any number of factors. Therefore, the conclusions of Theorem 1 are valid whenever parameters appear only in complete sets, which are sets such that if any interaction parameters between a group of factors are included, then all possible such parameters are included. Theorem 2 may therefore be generalized as follows.

THEOREM 3. Under any canonical parameterization involving only complete sets, or a nonsingular transformation of such a parameterization, the determinant and average variance criteria are satisfied if and only if the design is orthogonal under the canonical parameterization.

It is of considerable interest to determine which designs satisfy the two criteria if the parameterization does not consist only of complete sets. The simplest such situation involves two factors, one at two and the other at three levels. The main effects for the three-level factor are the linear effect and the quadratic effect, as in the example discussed earlier. Similarly, the interaction between the two factors may be parameterized using linear and quadratic components. Of the two, suppose only the quadratic component is included in the model. Then the matrix Z is as follows:

$$Z = \begin{bmatrix} 1 & -1 & -\sqrt{3/2} & 1/\sqrt{2} & -1/\sqrt{2} \\ 1 & -1 & 0 & -2/\sqrt{2} & 2/\sqrt{2} \\ 1 & -1 & \sqrt{3/2} & 1/\sqrt{2} & -1/\sqrt{2} \\ 1 & 1 & -\sqrt{3/2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 1 & 1 & 0 & -2/\sqrt{2} & -2/\sqrt{2} \\ 1 & 1 & \sqrt{3/2} & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

The rows of Z correspond to treatment combinations 00, 01, 02, 10, 11, and 12, respectively. Consider designs for which the number of runs at treatment combinations 00, 02, 10, and 12 is a constant λ and the number of runs at combinations 01 and 11 is $\frac{1}{2}(N-4\lambda)$. For such designs the cross-product matrix is of the form

$$X'X = \begin{bmatrix} N & 0 & 0 & \sqrt{2}(6\lambda-N) & 0 \\ 0 & N & 0 & 0 & \sqrt{2}(6\lambda-N) \\ 0 & 0 & 6\lambda & 0 & 0 \\ \sqrt{2}(6\lambda-N) & 0 & 0 & 2N-6\lambda & 0 \\ 0 & \sqrt{2}(6\lambda-N) & 0 & 0 & 2N-6\lambda \end{bmatrix}.$$

We seek the value of λ/N which will minimize $N^5/\det(X'X) = N^5/3^5 2^3 \lambda^3 (N-4\lambda)^2$. This value is $\lambda/N = 3/20$, for which the criterion has approximately the value .95 (the value for the orthogonal design being 1.0). The trace of the inverse times N is $(3N^2-8\lambda N)/6\lambda(N-4\lambda)$, for which the minimizing value is $\lambda/N = (3-\sqrt{3})/8$. The criterion has the value 4.98, again slightly better than the value of 5.00 obtained with the orthogonal design.

REFERENCE

[1] Plackett, R. L, and Burman, J. P., "The design of optimum multifactorial experiments," Biometrika, 33 (1946), 305-325.

APPENDIX D

DESIGNS FOR STUDYING ONE FACTOR AT A TIME

INTRODUCTION

It has become axiomatic in the statistical experimental design literature to discourage the practice of varying one factor at a time. In the case of factors each at two levels, the support for this point of view is that the variances of the main-effect estimates using such designs are considerably larger than with orthogonal designs.

On the other hand, the experimenter often likes such designs because he finds out more rapidly whether a new factor has any effect. He continually receives information rather than having to wait till the entire experiment is completed. If the magnitudes of the effects he is interested in are several times as large as experimental error, if he does not need to describe these effects precisely, and if there are no interactions, there is no particular disadvantage in experimenting in this way.

Another advantage of one-at-a-time experiments is that they are contractible or expansible without limit. Thus, no matter how many tests have been run, the experiment may be stopped, in which case estimates of the effects of those factors which have already been varied may be obtained, or it may be continued by the introduction of a new two-level factor.

An excellent exposition of the statistical arguments against such designs was given by Fisher in The Design of Experiments, sections 37 and 38 [2]. He bases his attack on their low efficiency compared with orthogonal designs and their lack of information about interactions.

Cuthbert Daniel [1] has presented the positive aspects of these designs. He pointed out that they are expansible and contractible and that they provide a quick look at each factor. He went on to say that they can often be augmented to form a half replicate plus one additional run, in which case the lost efficiency is for the most part regained.

It is the purpose of this paper to investigate the properties of one-at-a-time designs. It is pointed out that there are many classes of such designs. A lower bound is obtained for the variances of the estimates using any such design, and several classes are given for which the lower bound is obtained. The assumptions are that each factor has only two levels, that no interactions are present, and that there is no effect of the order of observations on the response. The last assumption is necessary because the order of the runs cannot be randomized. The number of factors need not be specified in advance.

BOUNDS FOR THE VARIANCES OF THE EFFECTS

We consider factors at two levels. For convenience, the initial level of each factor will be considered the low level and will be denoted by 0 or -. The first run will have each factor at its low level. If there are n factors the design contains $n + 1$ runs. The i-th factor appears for the first time at its high level in run $i + 1$. After it has been introduced it may stay at its high level, revert to its low level, or be varied between its two levels on subsequent tests. Thus there is a wide latitude of possible one-at-a-time designs.

Experiments for estimating the main effects of two-level factors are conventionally analyzed in terms of the coefficient matrix X as follows. The first column of X corresponds to the grand mean and has all its components equal to 1. Each of the remaining columns corresponds to one of the factors, and each row corresponds to a run. According to whether a given factor is at its high or low level in a given run, the corresponding element of X contains the entry $+1$ or -1.

Suppose a vector Y of N responses is obtained from the experiment. Under the assumption that there are no interactions we may write $Y = X\beta + e$, where β is the vector of the unknown parameters and e is a vector of independent random errors having mean zero and common variance σ^2. The least-squares estimate $\hat{\beta}$ of β is $\hat{\beta} = (X'X)^{-1}X'Y$.

The covariance matrix of $\hat{\beta}$ is $\sigma^2(X'X)^{-1}$.

In addition to working with the traditional coefficient matrix X, it will be convenient to introduce a reduced matrix R. Where X has an element 1, R also has 1; where X has a -1, R has a zero. It may be verified that X and R are related through the triangular transformation matrix T according to the equation X = RT, as in the following example.

$$\begin{bmatrix} 1 & -1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

In general, t_{ij} is given by the following rules:

$$t_{11} = 1$$

$$t_{1j} = -1 \quad (j=2, \ldots, n+1)$$

$$t_{jj} = 2 \quad (j=2, \ldots, n+1)$$

$$t_{ij} = 0 \quad \text{(otherwise)}.$$

It may be verified that

$$TT^{-1} = \begin{bmatrix} 1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} = I. \qquad (2)$$

If X is the coefficient matrix for any one-at-a-time design, then the corresponding R is lower triangular, and has 1's down the main diagonal. Therefore the determinant of R is unity. Since the elements of R are integers any minor is integral. Since the elements of an inverse are by definition an appropriate minor divided by the determinant of the original matrix, R^{-1} also consists of integers. It follows from (2) that each element of $X^{-1} = T^{-1}R^{-1}$ must be a multiple of $\frac{1}{2}$.

THEOREM 1. For a one-at-a-time design containing n factors at two levels and $n + 1$ runs, a lower bound for the variance of any estimate is $\frac{1}{2}\sigma^2$.

PROOF. The variances of the estimates are σ^2 times the diagonal elements of $(X'X)^{-1}$. Because X is square, $(X'X)^{-1}$ reduces to $(X^{-1})(X^{-1})'$. The diagonal elements are therefore the sums of squares of the elements in each row of X^{-1}. We know already that the elements of X^{-1} are all multiples of $\frac{1}{2}$. The sum of squares of the elements in a row must therefore

be a positive multiple of $\frac{1}{4}$. If the value were $\frac{1}{4}$, then all the elements would be zero except one which was equal to $\pm\frac{1}{2}$. The inner product of a row of X^{-1} and a column of X must of course be either 0 or 1. Since the elements of X are all either +1 or -1, the inner product of any column of X with a row containing a single $\frac{1}{2}$ would be $\pm\frac{1}{2}$. Therefore a lower bound to the sum of squares of elements of any row of X^{-1} is $\frac{1}{2}$, and the theorem is proved.

CONSTRUCTION AND COMPARISON OF VARIOUS ONE-AT-A-TIME DESIGNS

The most familiar family of one-at-a-time designs are those in which each factor returns to its low level after it has been "studied". The general form of the matrices R, R^{-1}, and X^{-1} in this family are exemplified by the following five-factor case:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} ; \quad R^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} ; \quad X^{-1} = \begin{bmatrix} -1\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} .$$

It will be noted that the variance of each main-effect estimates is $\frac{1}{2}\sigma^2$, the theoretical lower bound. The variance of the grand mean is $\sigma^2(n^2-3n+4)/4$ for the n-factor case.

It is of interest to inquire whether or not there is a family of designs in which the variance of the grand mean is also at the minimum level of $\frac{1}{2}\sigma^2$. The family of designs in which each factor is maintained at its high level satisfies this requirement. Again using a five-factor example to illustrate the general case, R, R^{-1}, and X^{-1} are as follows:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad ; \quad R^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad ; \quad X^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

These two classes of designs are two extremes in which all factors are either returned to their initial level or are held at their new level. Ordinarily an experimenter would like to determine which level of each factor is better, and conduct the remaining experiments at the more desirable level. Thus, in general some factors will be held at their new level and some will be returned to their old level. In a sense, the experiment will be midway between the extremes already investigated. For these extremes, the variances of the main effects attain the lower bound of $\frac{1}{2}\sigma^2$. The fact that the grand mean has a higher variance with the first family considered will often make little difference in the type of experiment in which one-at-a-time designs are often used. It seems appropriate to ask whether or not the main-effect variances will always be $\frac{1}{2}\sigma^2$ if factors are maintained at their more desirable level for the duration of the experiment. The answer is yes as can be seen from the following theorem.

THEOREM 2. For any one-at-a-time design in which each factor is maintained at either the high or low level after the initial introduction of the high level, the variances of the main-effect estimates are all $\frac{1}{2}\sigma^2$.

PROOF. Except for the first row, each entry in the matrix X^{-1} is $\frac{1}{2}$ times the corresponding entry in R^{-1}. Therefore the theorem will be proved by showing that each row of R^{-1} except the first contains one +1, one -1, and the remaining entries are 0. The proof will be inductive.

The theorem is obviously true for the case of one factor. Two possibilities exist for $n = 2$:

$$R^{(1)} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}; \quad R^{(1)-1} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}; \quad R^{(2)} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}; \quad R^{(2)-1} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

In either case R^{-1} has the required form.

Assume the theorem is true for designs involving n or fewer factors. The reduced design matrix R_2 for a design involving $n + 1$ factors may be partitioned as follows:

$$R_2 = \left[\begin{array}{c|c|c} R_o & z & z \\ \hline r & 1 & 0 \\ \hline r & q & 1 \end{array}\right]$$

Here R_o is the $n \times n$ reduced design matrix for a design of the required type involving $n - 1$ factors, and z is an n-component vector of zeros. The vector r consists of zeros and ones, and is repeated in the last two rows because of the restriction that a factor is held at either its high or low level after it has initially been varied. The single entry q will either be zero or one depending on whether the n-th factor is to be maintained at its high or low level for the duration of the experiment. The inverse of R_2 may be partitioned in the same way:

$$R_2^{-1} = \left[\begin{array}{c|c|c} R_o^{-1} & z & z \\ \hline a & 1 & 0 \\ \hline b & c & 1 \end{array}\right] .$$

Since R_2 is a lower triangular matrix with R_1 as a submatrix, R_2^{-1} is also lower triangular with R_1^{-1} (and hence also R_o^{-1}) as a corresponding submatrix. Thus, the first n rows of R_2^{-1} are as indicated above, where z again indicates an n-component zero vector. By the induction hypothesis, the n-component vector a has all its entries zero except one, which is equal to -1. Suppose the component q in R_2 is equal to 1. Then in order for the inner products of the last row of R_2 and the first $n + 1$ columns of R_2^{-1} all to have the value zero, the vector b must be an n-component zero (row) vector and c must have the value -1. Now suppose q is equal to zero. In order

for the inner product of the last row of R_2 and the first $n+1$ columns of X_2 all to have the value zero, b must equal a and c must equal zero. In either case, the last row of R_2^{-1} contains one 1, one -1, and n 0's. The theorem now follows by induction.

This theorem characterizes the "optimum class" of one-at-a-time designs. We conclude by giving an example of a very bad one-at-a-time design which is not a member of this class. Suppose instead of returning each factor to its initial low level immediately, it is kept at the high level for one additional run. Intuitively it would seem that the latter series of designs would differ little from the former, except that the high level is "studied" a little further. In fact this minor modification seriously degrades the quality of the designs. A five-factor example of this series is as follows:

$$R = \begin{bmatrix} 1&0&0&0&0&0 \\ 1&1&0&0&0&0 \\ 1&1&1&0&0&0 \\ 1&0&1&1&0&0 \\ 1&0&0&1&1&0 \\ 1&0&0&0&1&1 \end{bmatrix}; \quad R^{-1} = \begin{bmatrix} 1&0&0&0&0&0 \\ -1&1&0&0&0&0 \\ 0&-1&1&0&0&0 \\ -1&1&-1&1&0&0 \\ 0&-1&1&-1&1&0 \\ -1&1&-1&1&-1&1 \end{bmatrix}; \quad X^{-1} = \begin{bmatrix} -\frac{1}{2}&\frac{1}{2}&0&\frac{1}{2}&0&\frac{1}{2} \\ -\frac{1}{2}&\frac{1}{2}&0&0&0&0 \\ 0&-\frac{1}{2}&\frac{1}{2}&0&0&0 \\ -\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&0&0 \\ 0&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&0 \\ -\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2}&\frac{1}{2} \end{bmatrix}.$$

In general the first two main-effect estimates will have variance $\frac{1}{2}\sigma^2$, the next two σ^2, the next two $1\frac{1}{2}\sigma^2$, etc. This series of design illustrates the fact that apparently inconsequential modifications of designs can in fact have a serious effect on the variances of the estimates.

REFERENCES

[1] Daniel, C., (1958). On Varying One Factor At A Time, Biometrics, 14, 430–431.

[2] Fisher, R. A. (1942). The Design of Experiments, (3rd ed.). Oliver and Boyd, Edinburgh.

APPENDIX E

SOME NEW INCOMPLETE FACTORIAL DESIGNS

SUMMARY

An exhaustive search of possible twelve-run 2^4 designs has been made. There are nineteen essentially different design configurations, fifteen of which permit the estimation of the mean, main effects, and two-factor interactions. Two have been previously discussed in the literature. Two new designs seem to be of practical value. One has variances of $(15/128)\sigma^2$ for main effects and $(1/8)\sigma^2$ for two-factor interactions; the other has variances of $(7/64)\sigma^2$ for the main effects, but the interactions have somewhat higher variances.

An incomplete 2^7 design with 29 runs has been found in which all main effects and interactions can be estimated with variance $.050\sigma^2$. A modification with 36 runs allows estimates of the main effects with variance $.039\sigma^2$ and of two-factor interactions with variance $.048\sigma^2$.

INTRODUCTION

This paper was originally given in November of 1961 at a Central Regional Meeting of the Institute of Mathematical Statistics held in Urbana, Illinois. The research work was done at the University of Chicago while the author was a graduate student in the Department of Statistics. The

paper describes two loosely associated investigations on 2^n incomplete factorial designs which were not, however, incorporated into the author's thesis. They are of some general interest, and have been referred to in a general paper on non-orthogonal designs [2] by the author. Since the latter paper has been submitted for publication, and with the anticipation that its publication will create some interest in these results, this paper has been prepared at this time.

FOUR-FACTOR TWELVE-RUN DESIGNS

In the case of four two-level factors, there are four main effects, six two-factor interactions, four three-factor interactions, and one four-factor interaction. If interactions involving more than two factors are negligible, there are, including the mean, eleven parameters to be estimated. Two twelve-run designs have been proposed by Peter John [1] to handle such situations, although he has given no indication as to whether or not his designs are in any sense optimum. In an attempt to determine the relative merits of his designs, I have made an exhaustive classification of all possible four-factor designs containing runs at twelve distinct treatment combinations. It is most convenient to perform this classification in terms of the four treatment combinations which are omitted.

As usual, A, B, C, and D denote the factors. The treatment combination with every factor at its low level will be denoted by (1), and every other treatment combination by those (lower case) letters corresponding to the factors at their high levels. The treatment combinations form a group whose identity is (1). The product of two treatment combinations is the treatment combination containing all the letters from the original two except any they have in common. Thus, for example, $a \times b = ab$, $abc \times cd = abd$, and $ab \times ab = (1)$.

There are $\binom{16}{4} = 1820$ possible ways of choosing four treatment combinations to be omitted, but many of these choices lead to equivalent designs. Two designs will be considered equivalent if one can be obtained from the other by a combination of one or both of the following two operations: 1) rearranging the factors, and 2) reversing, for any factors, the high and the low levels. It follows that we may, without loss of generality, restrict attention to only those quadruples which contain any given fixed treatment combination; the treatment combination (1) will be so chosen. The 39 distinctive sets of four treatment combinations given in Table I all contain the identity (1) and are such that one cannot be obtained from another by rearranging the factors. Thus, given any set of four treatment combinations containing (1), by relabeling the variables one of the sets in Table I will be obtained. For example, [(1), c, abd, ac] becomes [(1), a, ab, bcd] which is

TABLE I

DISTINCTIVE FOUR-POINT CONFIGURATIONS

(1),a,b,c

(1),a,b,ab

(1),a,b,ac

(1),a,b,cd

(1),a,b,abc

(1),a,b,acd

(1),a,b,abcd

(1),a,ab,ac = [(1),a,b,c]

(1),a,ab,bc = [(1),a,b,abc]

(1),a,ab,cd = [(1),a,b,acd]

(1),a,bc,bd

(1),a,ab,abc = [(1),a,b,ac]

(1),a,ab,acd = [(1),a,b,cd]

(1),a,ab,bcd = [(1),a,b,abcd]

(1),a,bc,abc

(1),a,bc,abd

(1),a,bc,bcd

(1),a,ab,abcd = [(1),a,b,acd]

(1),a,bc,abcd

(1),a,abc,abd = [(1),a,bc,bd]

(1),a,abc,bcd = [(1),a,bc,abcd]

(1),a,abc,abcd = [(1),a,bc,bcd]

(1),a,bcd,abcd

(1),ab,ac,ad

(1),ab,ac,bc

(1),ab,ac,bd

(1),ab,ac,abc = [(1),a,b,abc]

(1),ab,ac,abd = [(1),a,bc,bd]

(1),ab,ac,bcd

(1),ab,cd,abc = [(1),a,bc,abcd]

(1),ab,ac,abcd = [(1),ab,ac,bd]

(1),ab,cd,abcd

(1),ab,abc,abd = [(1),a,b,cd]

(1),ab,abc,acd = [(1),a,bc,abd]

(1),ab,acd,bcd

(1),ab,abc,abcd = [(1),a,b,acd]

(1),ab,acd,abcd = [(1),a,bc,abcd]

(1),abc,abd,acd = [(1),ab,ac,bcd]

(1),abc,abd,abcd = [(1),a,b,abcd]

the 14th entry. The process of interchanging the levels of one or more factors is equivalent to multiplying each combination in the set of four by another treatment combination, namely that containing the letters for the factors whose levels are changed. Since by convention each set is required to contain the identity (1), all possible equivalent sets can be obtained by multiplying the given set by each of its elements in turn. Thus, starting with the set [(1),a,b,acd], three equivalent sets are obtained as follows:

TABLE II

Multiplier	Derived Set	Standard Form
(1)	(1),a,b,acd	(1),a,b,acd
a	a,(1),ab,cd	(1),a,ab,cd
b	b,ab,(1),abcd	(1),a,ab,abcd
acd	acd,cd,abcd,(1)	(1),ab,abc,abcd

The final column, labeled "Standard Form", is obtained from the second by renaming and rearranging the factors so that the sets are in the same form as in Table I. In Table I those sets which are equivalent to sets preceding them in the table have the equivalent set given after them in square brackets.

By examination of Table I it can be seen that there are exactly nineteen equivalence classes to which the possible twelve-run designs belong. For those designs which are nonsingular, the variances of the estimates using these designs are given in Table III. Each variance is actually a multiple of σ^2, and only the coefficients of σ^2 are given in the table. Those designs from which it is impossible to estimate all eleven parameters have the notation "singular" entered in the table.

The designs proposed by John are numbers 15 and 19 in Table III. Both have the property that the four treatment combinations omitted are subgroups, so that these designs are fractional factorials. The only other designs with this property are the four singular designs. Both of John's designs have the variance of each estimate equal to $\sigma^2/8$ except for two estimates which have variance $3\sigma^2/32$. These are the grand mean and a main effect for 15 and the grand mean and an interaction for 19. Since main effects are ordinarily of more interest, 15 would often be preferable to 19, as John points out.

Three other designs appear to have merit; these are numbers 14, 16, and 17. The average variance of the estimates is as low as possible for designs 14, 15, 17, and 19. Design number 14 is for most purposes the best twelve-run design available, since it has the added advantage that each main effect is estimated with a variance which is below this

TABLE III

VARIANCES OF ESTIMATES FOR 12-RUN 2^4 DESIGNS

Number	Combinations Omitted	Variances (Times 128)											
		Mean	A	B	C	D	AB	AC	AD	BC	BD	CD	Det X'X
1	(1), a, b, c	48	28	28	28	48	16	16	28	16	28	28	$.34 \times 10^{11}$
2	(1), a, b, ab	singular											0
3	(1), a, b, ac	32	28	28	28	32	16	16	28	16	28	28	$.34 \times 10^{11}$
4	(1), a, b, cd	32	16	16	60	60	16	28	28	28	28	32	$.34 \times 10^{11}$
5	(1), a, b, abc	32	28	28	60	32	16	16	28	16	28	60	$.34 \times 10^{11}$
6	(1), a, b, acd	32	16	16	28	28	16	28	28	28	28	32	$.34 \times 10^{11}$
7	(1), a, b, abcd	48	16	16	28	28	16	28	28	28	28	48	$.34 \times 10^{11}$
8	(1), a, bc, bd	16	14	32	24	24	14	14	14	24	32	24	$.69 \times 10^{11}$
9	(1), a, bc, abc	singular											0
10	(1), a, bc, abd	16	14	32	24	24	14	14	14	24	24	32	$.69 \times 10^{11}$
11	(1), a, bc, bcd	32	16	28	28	16	28	28	16	32	28	28	$.34 \times 10^{11}$
12	(1), a, bc, abcd	32	16	28	28	16	28	28	16	32	60	60	$.34 \times 10^{11}$
13	(1), a, bcd, abcd	singular											0
14	(1), ab, ac, ad	12	15	15	15	15	16	16	16	16	16	16	$.14 \times 10^{12}$
15	(1), ab, ac, bc	12	16	16	16	12	16	16	16	16	16	16	$.14 \times 10^{12}$
16	(1), ab, ac, bd	16	14	14	14	14	32	24	24	24	24	32	$.69 \times 10^{11}$
17	(1), ab, ac, bcd	12	16	16	16	15	16	16	15	16	15	15	$.14 \times 10^{12}$
18	(1), ab, cd, abcd	singular											0
19	(1), ab, acd, bcd	12	16	16	16	16	16	16	16	16	16	12	$.14 \times 10^{12}$

average. Design 16 would be useful if the interactions were of only incidental interest, since the main effects have the lowest average variance with this design. Design 17 would be useful if one factor (D) and its interactions were of the most interest.

It should be noted that design 14 is equivalent to a permutation-invariant design. The concept of permutation-invariance, which I introduced in another paper [2], implies that all factors are treated alike in the sense that the cross-product matrix remains unaltered if the factors are permuted. By multiplying each treatment combination by a the combinations omitted are seen to be a, b, c, and d, and this set can be seen to be permutation-invariant. Design 14 is the only design of the nineteen which is equivalent to a permutation-invariant design.

NEW PERMUTATION-INVARIANT RESOLUTION-FIVE DESIGNS

The class of fractional factorials is deficient for some experimental situations, one of which is the case of the 2^7 when interactions are of interest. Although only 29 estimates are to be obtained, the smallest fractional factorial is the half-replicate, containing 64 runs. This fact is especially disappointing when one recalls that 8 factors can be accommodated in the same number of runs. Therefore there is considerable

interest in finding an incomplete factorial from which all the effects and interactions can be estimated efficiently and which is considerably smaller than the smallest fractional factorial available. Two such designs, both permutation-invariant, have been found for this experimental situation. They were derived by analogy with the permutation-invariant 12-run 2^4 design described above.

The first of the new 2^7 designs we will consider contains just 29 runs, the smallest possible number. These runs are as follows: i) one run at the treatment combination (1); ii) one run at each of the 21 two-letter treatment combinations (i.e., ab, ac, $\cdots$, fg); and iii) one run at each of the 7 six-letter treatment combinations (i.e., abcdef, $\cdots$, bcdefg). The estimate of the grand mean is $-(1/6)Y_{(1)} + (1/24)[Y_{ab} + \cdots + Y_{bcdefg}]$. Here $Y_{(1)}$, Y_{ab}, etc., denote the response at treatment combinations (1), ab, etc. Its variance is $.076\sigma^2$. The estimate of the main effect of factor A is $\hat{A} = -(1/24)Y_{(1)} + (1/24)[Y_{ab} + \cdots + Y_{ag}] - (1/48)[Y_{bc} + \cdots + Y_{fg}] + (1/24)[Y_{abcdef} + \cdots + Y_{acdefg}] - (7/48)Y_{bcdefg}$. The estimates of the other main effects are analogous. The variances of the main-effect estimates are each $.050\sigma^2$. The estimate of AB is

$$AB = (1/24)Y_{(1)} + (7/48)Y_{ab} - (1/24)[Y_{ac} + \cdots + Y_{bg}] + (1/48)[Y_{cd} + \cdots + Y_{fg}] - (1/24)[Y_{acdefg} + Y_{bcdefg}] + (1/48)[Y_{abcdef} + \cdots + Y_{abdefg}].$$

The other interactions are defined analogously. The variance of these estimates is also $.050\sigma^2$.

In comparing incomplete factorial designs it is useful to define the efficiency of a design for a given estimate as the ratio of the per-run information with the given design to the per-run information with the full factorial, where the information is the reciprocal of the variance. For this design the efficiencies are .45 for the grand mean and .68 for the main effects and interactions.

The second of the new 2^7 designs contains i) one run at (1); ii) one run at the 21 two-letter treatment combinations; and iii) <u>two</u> runs at each of the 7 six-letter combinations. Thus, the second design can be obtained from the first by duplicating each of the six-letter treatment combinations. This 36-run design has an efficiency of .40 for the mean, .81 for the main effects, and .58 for the two-factor interactions. The high efficiency of the main-effect estimates and the fact that the design contains seven duplicated points from which an estimate of pure experimental error can be obtained, make this a particularly attractive design.

A six-factor design analogous to the first seven-factor design has also been investigated. This design contains 22 runs as follows: i) one at (1); ii) one at each two-letter combination; and iii) one at each five-letter combination. The efficiencies are quite high, namely .83 for the mean and .87 for the main effects and interactions.

REFERENCES

[1] JOHN, P. W. M. (1961). Three-quarter replicates of 2^4 and 2^5 designs. Biometrics 17 319-321.

[2] WEBB, S. R. (1964). Characterization of non-orthogonal incomplete factorial designs. Research Report RR 64-18, Rocketdyne, Canoga Park. (Appendix B of the present report).